

PHILOSOPHY OF SCIENCE

A Unified Approach

GERHARD SCHURZ

Second Edition



PHILOSOPHY OF SCIENCE

A Unified Approach

GERHARD SCHURZ

Second Edition



“I consider this book to be a unique work both with regard to the sheer amount of material it covers but also with regard to the clarity, depth and insight with which these matters are dealt with by one of Europe’s foremost analytic philosophers. I would not hesitate to use it in an advanced course on the subject, and I believe it would work well also for an introductory course given the careful way in which the book is structured into introductory and advanced material.”

– **Erik J. Olsson** , *Lund University*

“ *Philosophy of Science: A Unified Approach* revamps and bolsters the logical empiricist way to think about and to practice philosophy of science. It combines technical precision and rigor with philosophical breadth and depth. It delves into all major philosophical problems and issues in general philosophy of science. It elegantly presents the received positions and approaches, but it also articulates novel views and arguments. It systematically explains the necessary technical concepts and makes them do deep philosophical work. *Philosophy of Science: A Unified Approach* lives up to its promise to highlight and defend the methodological unity of the sciences. Gerhard Schurz has written a superb introduction to philosophy of science; novices and experts alike have a lot to learn from it.”

Stathis Psillos , *University of Western Ontario and University of Athens*

“This is an excellent introduction to the general philosophy of science. It is lucidly written and very well structured in elementary and advanced subsections. Inspired by the classic authors of the 20th century, such as Carnap and Popper, and guided by the methodological version of the idea of the unity of the sciences and the humanities, Gerhard Schurz shows how it stands with the main topics in the first decades of the 21st century.”

– **Theo Kuipers** , *University of Groningen*

Philosophy of Science

Philosophy of Science: A Unified Approach, Second Edition, combines a general introduction to philosophy of science with an integrated survey of all its important subfields. As this book's subtitle suggests, this excellent overview is guided methodologically by "a unified approach" to philosophy of science. Behind the diversity of scientific fields, one can recognize a methodological unity. This unity is worked out in this book, revealing all the while important differences between subject areas.

This book offers a two-part approach that makes it an excellent introduction for readers new to the field and a useful resource for more advanced students. Each chapter is divided into two sections. The first section assumes no foreknowledge of the subject introduced, and the second section builds upon the first by bringing into the conversation more advanced, complementary topics from contemporary research.

Definitions, key propositions, examples, and figures complement all of the core material. At the end of every chapter, there are selected readings and exercise questions (with solutions and answers to many of them at the end of this book). This book also includes a comprehensive bibliography and an index.

The Second Edition has been revised to make the technical material much more accessible for students outside of the sciences and philosophy. It also incorporates new research over the last decade, including significantly expanded coverage of values in science, mechanistic explanations, and inductive risk. An extended index of subjects that can also be used as a glossary has been added to the back of this book.

Gerhard Schurz (MSc 1980, PhD 1983) is Professor of Philosophy at the Heinrich Heine University Düsseldorf (since March 2002) and member of the Leopoldina (German Academy of Sciences) and the International Academy for Philosophy of Science. Before 2002, he was an Associated Professor at the University of Salzburg and a Guest Professor at the University of California at Irvine and Yale University. His research areas include philosophy of science, epistemology, logic, and cognitive science. He is the author of 270 research articles and 12 books.

Philosophy of Science

A Unified Approach

Second Edition

Gerhard Schurz



Designed cover image: Getty Images

Second edition published 2026
by Routledge
605 Third Avenue, New York, NY 10158

and by Routledge
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2026 Gerhard Schurz

The right of Gerhard Schurz to be identified as author of this work has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

The Open Access version of this book, available at www.taylorfrancis.com, has been made available under a Creative Commons Attribution-ShareAlike 4.0 (CC BY-SA) International license.

Any third party material in this book is not included in the OA Creative Commons license, unless indicated otherwise in a credit line to the material. Please direct any permissions enquiries to the original rightsholder.

For Product Safety Concerns and Information please contact our EU representative
GPSR@taylorandfrancis.com. Taylor & Francis Verlag GmbH, Kaufingerstraße 24, 80331 München, Germany.

Trademark notice : Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

First edition published by Routledge 2013

Library of Congress Cataloging-in-Publication Data

Names: Schurz, Gerhard, 1956– author

Title: Philosophy of science : a unified approach / Gerhard Schurz.

Other titles: Einführung in die Wissenschaftstheorie. English

Description: 2nd edition. | New York, NY ; Abingdon, Oxon : Routledge, 2026. | Includes bibliographical references and index. | Text in English. Translation from German.

Identifiers: LCCN 2026005541 | ISBN 9781041145820 hardback | ISBN 9781041145813 paperback | ISBN 9781003675075 ebook

Subjects: LCSH: Science—Philosophy

Classification: LCC Q175 .S413414 2026

LC record available at <https://lcn.loc.gov/2026005541>

ISBN: 978-1-041-14582-0 (hbk)

ISBN: 978-1-041-14581-3 (pbk)

ISBN: 978-1-003-67507-5 (ebk)

DOI: [10.4324/9781003675075](https://doi.org/10.4324/9781003675075)

Typeset in Sabon
by codeMantra

Contents

[List of Figures, Definitions, and Key Propositions](#)

[List of Symbols](#)

[Foreword](#)

[Foreword to the Second Edition](#)

[1 Introduction: Where Do We Stand?](#)

- [1.1 Tasks and Aims of Philosophy of Science](#)
- [1.2 On the Historical Development of Philosophy of Science](#)
- [1.3 Philosophical Positions in Contemporary Philosophy of Science](#)
 - [1.3.1 Logical Empiricism](#)
 - [1.3.2 Critical Rationalism](#)
 - [1.3.3 The Historical Account and Relativism](#)
 - [1.3.4 Metaphysical and Scientific Realism](#)
 - [1.3.5 Naturalism and the Cognitive Turn](#)
 - [1.3.6 Structuralist and Semantic View of Theories](#)
 - [1.3.7 Pragmatic Philosophy of Science](#)
 - [1.3.8 Hermeneutics and Critical Theory](#)
 - [1.3.9 Radical Constructivism, Postmodernism, and the Science Wars](#)
- [1.4 Introductory Literature and Exercises](#)

[2 The Question of Unity: In Search for Common Foundations of the Sciences](#)

- [2.1 Normative or Descriptive? The Method of Rational Reconstruction](#)
- [2.2 Common Epistemological Assumptions](#)
- [2.3 Common Methodological Features](#)
- [2.4 Scientific Disciplines and Their Classification](#)
 - [2.4.1 The Special Status of Mathematical Sciences](#)
 - [2.4.2 Testing the Common Foundations of Science](#)
 - [2.4.3 Two Further Classifications of Sciences](#)
- [2.5 The Question of Value Neutrality](#)
 - [2.5.1 Max Weber's Value-Freedom Postulate](#)
 - [2.5.2 Value Neutrality and Means-End Inferences](#)
 - [2.5.3 Explication of the Value-Neutrality Requirement](#)
- [2.6 The Demarcation Problem and the Methodological Unity of Sciences](#)
 - [2.6.1 Demarcation](#)

2.6.2 Methodological Unity instead of Physical Reductionism

2.7 Scientific Inference and Argumentation

2.7.1 Deduction and Induction

2.7.2 Popper and the Significance of Inductive Inference in the Sciences

2.7.3 Abduction and Inference to the Best Explanation

2.7.4 Monotonic and Non-monotonic Inferences

Complementary and Advanced Topics

2.8 How Much Metaphysics is Presupposed by Science?

2.8.1 Truth, Idealism, and Realism

2.8.2 Constructive Realism versus Radical Constructivism

2.9 On the Theory (In)dependence of Observation

2.9.1 Theory-Dependency of Observation – Pros and Cons

2.9.2 A Definition of Theory-Neutral Observation Concepts

2.10 Value Neutrality in the Context of Contemporary Debates

2.10.1 The Role of Values in Science

2.10.2 The Argument of Inductive Risk

2.10.3 Meta-ethical Justification of Value Neutrality and the Difference between Experience and Valuation

2.11 The Justification of Induction: An Unsolvable Problem?

2.11.1 Is Induction Rationally Unjustifiable?

2.11.2 The Optimality of Meta-induction

2.12 Introductory Literature and Exercises

3 The Conceptual Toolkit: Language, Logic, and Probability

3.1 Kinds and Classification of Concepts

3.1.1 The Logical Type of Concepts

3.1.2 Syntax and Semantics, Statements, and Models

3.1.3 The Content Type of Concepts

3.1.4 The Gradation (Scale) Type of Concepts

3.2 Classification of Sentences According to Their Content Type

3.2.1 Descriptive versus Prescriptive Sentences

3.2.2 Observation and Basic Sentences, Empirical and Theoretical Sentences

3.3 Logical Truth and Deductive Logic

3.4 Conventions of Meaning and Definitional Truths

3.5 Classification of Sentences According to Their Generality

3.6 General Sentences, Lawlikeness, Determinism, and Indeterminism

3.7 Content of Sentences and Kinds of Content

3.8 Verification, Falsification, Confirmation, and Disconfirmation

3.9 Objective (Statistical) and Subjective (Epistemic) Probability

Complementary and Advanced Topics

3.10 Disposition Concepts

3.11 Challenges to the Logical-Definitional-Synthetic Distinction

- [3.12 Relevance and Irrelevance in Logical Inferences](#)
- [3.13 More on Probability](#)
 - [3.13.1 Mathematical Laws of Probability](#)
 - [3.13.2 Problems of Objective-Statistical Probability Concepts](#)
 - [3.13.3 Problems of Subjective Probability Concepts](#)
 - [3.13.4 Connections between Objective and Subjective Probabilities: Principal Principle and Narrowest Reference Class](#)
- [3.14 Introductory Literature and Exercises](#)

4 A Question of Fit: Law Hypotheses and Their Empirical Testing

- [4.1 The Condition of Relevance](#)
 - [4.1.1 Relevance of Strict Laws](#)
 - [4.1.2 Relevance of Statistical Laws](#)
- [4.2 Testing of Strict Law Hypotheses for Truth and Relevance](#)
 - [4.2.1 The Method of Agreement and Difference](#)
 - [4.2.2 Methodological Induction: How to Discover Strict Law Hypotheses](#)
- [4.3 Testing Statistical Laws](#)
 - [4.3.1 Testing for Likely Truth and Acceptance Intervals](#)
 - [4.3.2 Discovery of Statistical Hypotheses and Confidence Intervals](#)
 - [4.3.3 Test Statistics and Inference Statistics](#)
 - [4.3.4 Testing for Relevance and Significant Differences](#)
 - [4.3.5 Statistical Representativeness](#)
 - [4.3.6 Sources of Error in Statistical Methods](#)
 - [4.3.7 Applying Statistical Hypotheses to Individual Cases](#)
- [4.4 Correlation and Causality](#)
 - [4.4.1 Hidden Variables](#)
 - [4.4.2 Causal Direction](#)

Complementary and Advanced Topics

- [4.5 Ceteris Paribus Laws and Normic Laws](#)
 - [4.5.1 Comparative versus Exclusive CP Laws](#)
 - [4.5.2 Normic Laws and Their Evolution–Theoretic Foundation](#)
- [4.6 Probability Distributions for Continuous Variables](#)
- [4.7 From Classical Statistics to Bayesian Statistics](#)
 - [4.7.1 The Likelihood Intuition and Its Bayesianism Justification](#)
 - [4.7.2 Objective Bayesianism and the Principle of Indifference](#)
 - [4.7.3 Subjective Bayesianism and Convergence of Opinions](#)
- [4.8 Introductory Literature and Exercises](#)

5 Going beyond Experience: Theories and Their Empirical Evaluation

- [5.1 Theoretical Concepts and Multiple Laws of Correspondence](#)
- [5.2 The Example of Newtonian Physics](#)
- [5.3 The Structure of Scientific Theories](#)
 - [5.3.1 Components of Theories](#)
 - [5.3.2 Theory Nets, Intertheoretical Relations, and Kinds of Theoretical Concepts](#)

- 5.4 [Methodological Features of \(Good\) Scientific Theories](#)
 - 5.4.1 [System Character: Holism of Meaning, Content, and Theory Testing \(Duhem's Thesis\)](#)
 - 5.4.2 [Empirical Creativity, Global Unification, and Novel Predictions: Answers to Hempel's Dilemma](#)
- 5.5 [The Example of Piaget's Theory of Cognitive Development](#)
- 5.6 [The Dynamics of Scientific Theories](#)
 - 5.6.1 [Lakatos' Model of Theory Revision](#)
 - 5.6.2 [Theory Evaluation and Theory Progress](#)
- 5.7 [Instrumentalism and Realism: The Ontology of Scientific Theories](#)

[Complementary and Advanced Topics](#)

- 5.8 [Instrumentalism-Realism Controversies in Philosophy of Science](#)
 - 5.8.1 [The Ramsey Sentence of a Theory: An Instrumentalist Challenge](#)
 - 5.8.2 [The Realist Interpretation of the Ramsey Sentence](#)
 - 5.8.3 [The No-Miracles Argument](#)
 - 5.8.4 [The Pessimistic Meta-Induction](#)
 - 5.8.5 [The Empirical Underdetermination of Theories](#)
 - 5.8.6 [Intertheoretical Correspondence, Structural and Selective Realism](#)
 - 5.8.7 [Criteria for Empirical Significance](#)
 - 5.8.8 [Probabilistic Analysis of the NMA, Alternative Theories, and Local Realism](#)
- 5.9 [General Accounts of Confirmation](#)
 - 5.9.1 [Hypothetico-Deductive Confirmation](#)
 - 5.9.2 [Probabilistic Confirmation and Bayesianism](#)
 - 5.9.3 [Bayesian Pseudo-Confirmation through Content Cutting versus Genuine Confirmation](#)
 - 5.9.4 [Goodman's Paradox](#)
 - 5.9.5 [Confirmation and Acceptance](#)
- 5.10 [Testing Theories: Parameter Fitting and Use-Novel Evidence](#)
 - 5.10.1 [Post-Factum Speculations](#)
 - 5.10.2 [Curve Fitting](#)
- 5.11 [Non-Confirmational Accounts of Theory Evaluation](#)
 - 5.11.1 [Truthlikeness](#)
 - 5.11.2 [Unification, Coherence, and Simplicity](#)
- 5.12 [The Problem of Language Dependence](#)
- 5.13 [Introductory Literature and Exercises](#)

[6 In Search of Causes: Explanation and All That Goes with It](#)

- 6.1 [The Deductive-Nomological Model of Explanation](#)
- 6.2 [Explanation versus Prediction and Justification](#)
 - 6.2.1 [Predictions without Explanatory Function](#)
 - 6.2.2 [Causality and Lawlikeness](#)
 - 6.2.3 [Explanations without Predictive Function](#)
 - 6.2.4 [Logical Problems of Irrelevance and Redundancy](#)

6.3 Probabilistic Explanation Models

6.3.1 Inductive-Statistical Explanation

6.3.2 Maximal Specificity

6.3.3 Probabilistic and Causal Relevance Models

6.3.4 Conflicting Intuitions about the Magnitude of Explanatory Probabilities

6.4 Further Types of Explanation

6.4.1 Mechanistic Explanations

6.4.2 Normic Explanations and the Explanation of Human Actions

6.4.3 Explanation and Understanding as Unification

Complementary and Advanced Topics

6.5 Lawlikeness

6.5.1 Laws of Nature versus System Laws

6.5.2 Lawlikeness in the Wide versus Narrow Sense

6.5.3 Lawlikeness i.w.s. and Counterfactuals

6.5.4 Lawlikeness i.n.s., Spatiotemporal Universality, and Maxwell's Condition

6.5.5 The Best System Account

6.5.6 Physical Necessity and Independent Possibility Knowledge

6.6 Causality

6.6.1 Singular and General Causality

6.6.2 Regularity Accounts

6.6.3 Counterfactual Accounts

6.6.4 Causation and the Direction of Time

6.6.5 Causal Process Theories

6.6.6 Interventionist Accounts

6.6.7 Causality as a Theoretical Concept

6.6.8 The Theory of Causal Nets

6.6.9 Causal Discovery and Empirical Content of TCN

6.6.10 From Variables to Events: Overdetermination and Causal Preemption

6.7 Introductory Literature and Exercises

Bibliography

Solutions to Selected Exercises

Index of Names

Index of Subjects

List of Figures, Definitions, and Key Propositions

Figures

- [2.1-1 The method of rational reconstruction](#)
- [2.3-1 The three levels of the scientific method](#)
- [2.4-1 Classification of the factual sciences](#)
- [2.5-1 Schematic representation of the value neutrality requirement](#)
- [2.7-1 The inductive-deductive scheme](#)
- [2.7-2 Interaction of epistemic induction and abduction](#)
- [2.9-1 Duck-rabbit](#)
- [2.9-2 Mueller-Lyer illusion](#)
- [2.9-3 Expected learning curves in ostensive learning experiments](#)
- [3.1-1 Division of concepts according to their logical type](#)
- [3.1-2 Classification of concepts according to their content type](#)
- [3.1-3 Classification of concepts according to their gradation type](#)
- [3.1-4 Ordinal scale](#)
- [3.2-1 Classification of sentences according to their content type](#)
- [3.5-1 Classification of sentences according to their generality](#)
- [3.5-2 Statistical generalizations](#)
- [3.8-1 Testability of generalizations](#)
- [3.9-1 Non-monotonicity of conditional probabilities](#)
- [3.13-1 Three binomial distributions](#)
- [4.1-1 Irrelevance and joke](#)
- [4.3-1 Acceptance interval](#)
- [4.3-2 Acceptance and confidence interval](#)
- [4.3-3 Probability distribution of sample differences and significant difference](#)
- [4.4-1 Common cause](#)
- [4.4-2 Common cause vs. intermediate cause](#)
- [4.4-3 Spurious causal independence](#)
- [4.4-4 What is the direction of causality?](#)
- [4.6-1 Normal distribution](#)
- [4.7-1 Shortest interval](#)
- [4.7-2 Transformation of uniform into non-uniform density](#)
- [5.1-1 One theoretical feature causes many empirical dispositions](#)
- [5.1-2 Beam balances and spring scales on Earth and Uranus](#)
- [5.2-1 Theory lattice in classical mechanics](#)

[5.3-1,2,3 Subdivisions of a theory](#)
[5.3-4 Theory net of physical and chemical theories](#)
[5.6-1 Historical theory sequence](#)
[5.6-2 Theory comparison](#)
[5.6-3 Empirically complementary theories](#)
[5.6-4 Theories of optics after Bunge](#)
[5.10-1 Curve fitting and use-novel evidence](#)
[6.6-1 Explanation of screening-off](#)
[6.6-2 Three statistically strongly indistinguishable graphs](#)
[6.6-3 Causal graph of variables and related graphs of values](#)

Definitions

[\(Def. 2.1-1\) Method of rational reconstruction](#)
[\(Def. 2.5-1\) The general scheme of means-end inferences](#)
[\(Def. 2.7-1\) Monotonicity](#)
[\(Def. 2.9-1\) Theory-neutral observation concepts](#)
[\(Def. 3.2-1\) Descriptive versus prescriptive sentences](#)
[\(Def. 3.2-2\) Observational, empirical, theoretical and T-theoretical sentences](#)
[\(Def. 3.3-1\) Logical truth](#)
[\(Def. 3.3-2\) Logical validity](#)
[\(Def. 3.4-1\) Definitional truth](#)
[\(Def. 3.5-1\) Essential sentence types](#)
[\(Def. 3.7-1\) Logical and empirical content](#)
[\(Def. 3.8-1\) Verification, falsification, confirmation and disconfirmation](#)
[\(Def. 3.9-1\) Statistical and subjective probability](#)
[\(Def. 3.9-2\) Conditional probability](#)
[\(Def. 3.9-3\) Principle of narrowest reference class](#)
[\(Def. 3.10-1\) Bilateral reduction sentence](#)
[\(Def. 3.10-2\) Nomological definition of dispositions](#)
[\(Def. 3.12-1\) Relevance in logical inferences](#)
[\(Def. 3.12-2\) Content elements](#)
[\(Def. 3.12-3\) Relevant content](#)
[\(Def. 3.12-4\) Irreducible representation](#)
[\(Def. 4.1-1\) Condition of relevance for strict law hypotheses](#)
[\(Def. 4.1-2\) Statistical relevance for conjunctive antecedents](#)
[\(Def. 4.3-1\) Acceptance interval](#)
[\(Def. 4.3-2\) Significant difference](#)
[\(Def. 4.6-1\) Mean, variance and standard deviation](#)
[\(Def. 5.3-1\) Kinds of theoretical concepts](#)
[\(Def. 5.8-1\) Ramsey sentence](#)
[\(Def. 5.8-2\) Carnap sentence](#)
[\(Def. 5.8-3\) Ramsey-Lewis sentence](#)
[\(Def. 5.9-1\) Classical HD-confirmation](#)
[\(Def. 5.9-2\) Relevant HD-confirmation](#)
[\(Def. 5.9-3\) Incremental and absolute Bayesian confirmation](#)
[\(Def. 5.9-4\) Genuine confirmation \(full and partial\)](#)
[\(Def. 5.9-5\) Goodman's grue predicate](#)
[\(Def. 5.11-1\) Popper-truthlikeness](#)
[\(Def. 5.11-2\) Schurz/Weingartner-truthlikeness](#)

[\(Def. 6.5-1\) Spatiotemporal universality](#)
[\(Def. 6.6-2\) Counterfactual Causality according to Lewis](#)
[\(Def. 6.6-3\) Interventionist causality after Woodward](#)
[\(Def. 6.6-4\) Screening-off between variables](#)
[\(Def. 6.6-5\) Linking-up between variables](#)
[\(Def. 6.6-6\) Causal Markov axiom](#)
[\(Def. 6.6-7\) Faithfulness condition](#)
[\(Def. 6.6-8\) Minimality and productivity](#)
[\(Def. 6.6-9\) Cause-effect relation between values of variables](#)

Key Propositions

[\(Prop. 2.1-1\) The supreme epistemic goal of science](#)
[\(Prop. 2.5-1\) Value neutrality requirement](#)
[\(Prop. 2.7-1\) General pattern of abduction](#)
[\(Prop. 2.7-2\) Inference to the best explanation](#)
[\(Prop. 2.7-3\) Application rule for monotonic inferences](#)
[\(Prop. 2.7-4\) Application rule for non-monotonic inferences](#)
[\(Prop. 2.11-1\) Universal optimality of AW](#)
[\(Prop. 3.1-1\) Translation between logical and set-theoretical language](#)
[\(Prop. 3.1-2\) Linguistic and model-theoretic theory representation](#)
[\(Prop. 3.4-1\) Requirement of empirical non-creativity](#)
[\(Prop. 3.5-1\) Important logical relationships](#)
[\(Prop. 3.9-1\) Basic axioms of probability](#)
[\(Prop. 3.13-1\) Theorems of unconditional probability](#)
[\(Prop. 3.13-2\) Theorems of conditional probability](#)
[\(Prop. 3.13-3\) Binomial formula](#)
[\(Prop. 3.13-4\) Subjective probabilities as coherent fair betting quotients](#)
[\(Prop. 3.13-5\) Statistical principal principle](#)
[\(Prop. 3.13-6\) Weak induction principles for subjective probabilities](#)
[\(Prop. 4.1-1\) Statistical relevance and correlation for qualitative properties](#)
[\(Prop. 4.2-1\) Testing a strict law hypothesis](#)
[\(Prop. 4.3-1\) Testing a statistical law hypothesis](#)
[\(Prop. 4.4-1\) Reichenbach conditions for screening off](#)
[\(Prop. 4.5-1\) Comparative CP law](#)
[\(Prop. 4.5-2\) Exclusive CP assertion](#)
[\(Prop. 4.5-3\) Normic interpretation of eCP laws](#)
[\(Prop. 4.6-1\) Mean value, variance and standard deviation of the distribution of sample means](#)
[\(Prop. 4.7-1\) Principle of Bayes-statistics](#)
[\(Prop. 4.7-2\) Bayesian justification of the likelihood-intuition](#)
[\(Prop. 4.7-3\) Consequences of the indifference principle](#)
[\(Prop. 4.7-4\) Bayesian convergence results](#)
[\(Prop. 4.7-5\) Biased priors](#)
[\(Prop. 5.1-1\) Synthetic nature of laws of correspondence for theoretical concepts](#)
[\(Prop. 5.1-2,3\) Biased indicators](#)
[\(Prop. 5.4-1\) Logical background of the holism of testing](#)
[\(Prop. 5.4-2\) Empirical creativity](#)
[\(Prop. 5.6-1\) Falsification of theory version](#)
[\(Prop. 5.6-2,3\) Successes and failures of theory versions](#)
[\(Prop. 5.6-4\) Confirmation of theory versions](#)

[\(Prop. 5.6-5\) Dynamical truthlikeness](#)
[\(Prop. 5.6-6\) Theory comparison and theory progress](#)
[\(Prop. 5.6-7\) Rational acceptability of theory versions](#)
[\(Prop. 5.6-8\) Untested empirical excess content](#)
[\(Prop. 5.6-9\) Confirmation of theory cores](#)
[\(Prop. 5.6-10\) Preference between theory cores](#)
[\(Prop. 5.6-11\) Rational acceptability of theory cores](#)
[\(Prop. 5.8-1\) Newman's problem](#)
[\(Prop. 5.8-2\) Howson's analysis of the NMA](#)
[\(Prop. 5.9-1\) HD-confirmation and Bayesian confirmation](#)
[\(Prop. 5.9-2\) Impossibility results for rational acceptance](#)
[\(Prop. 5.10-1\) Necessary criteria for transfer of probability increase](#)
[\(Prop. 6.1-1\) DN-explanation of events](#)
[\(Prop. 6.1-2\) DN-explanation of laws](#)
[\(Prop. 6.2-1\) Lawlikeness and causality in DN-explanations](#)
[\(Prop. 6.2-2\) Predictive function](#)
[\(Prop. 6.2-3\) Logical relevance conditions for DN-explanations](#)
[\(Prop. 6.3-1\) Inductive-statistical explanation](#)
[\(Prop. 6.3-2\) Maximal specificity](#)
[\(Prop. 6.4-1\) Normic explanation](#)
[\(Prop. 6.4-2\) Unification condition](#)
[\(Prop. 6.5-1\) Physical necessity](#)
[\(Prop. 6.6-1\) Singular and general causality](#)
[\(Prop. 6.6-2\) Two versions of the causal Markov axiom](#)
[\(Prop. 6.6-3\) Equivalence of productivity and minimality](#)
[\(Prop. 6.6-4\) Causal discovery algorithm](#)
[\(Prop. 6.6-5\) Empirical content of TCN](#)
[\(Prop. 6.6-6\) Causal nets with time-indexed variables](#)

List of Symbols

$\neg$	not
$\wedge$	and
$\vee$	or
$\rightarrow$	material implication (if-then)
$\leftrightarrow$	equivalent (if-and-only-if)
$\leftrightarrow_{\text{def}}$	equivalent by definition
$\forall x$	for all x
$\exists x$	there exists an x
$\exists!x$	there exists exactly one x
$=$	identical
$\neq$	not identical
$=_{\text{def}}$	identical by definition
$\square$	necessary
$\mathbb{N}$	set of natural numbers
$\mathbb{R}$	set of real numbers
$\in$	is an element of

$\{a_1, \dots, a_n\}$	set of individuals
$(a_1, \dots, a_n)$	ordered sequence of individuals
$\{x: Fx\}$	set of all individuals which possess property F
$\subseteq$	proper or improper subset
$\subset$	proper subset
$\cap$	intersection
$\cup$	union
$-$	applied to sets set-theoretical complement
$\times$	Cartesian product between sets
$\cdot$	ordinary product between numbers
$/$	division (example $x \cdot y / z = x$ times y divided through z)
$\square$	empty set
$\sum_i \in_I x_i$ and $\prod_i \in_I x_i$	sum and product (respectively) over all x_i with i in I
D	domain
D^n	n -fold Cartesian product of D
I	interpretation function
$\mathcal{L}$	formal language
$M(\mathcal{L})$	set of models for language $\mathcal{L}$
$\leq$	smaller or equal
$\geq$	greater or equal
$\equiv$	equivalence relation

$f:A \rightarrow B$	function from domain A into set of values B
$\models$	verifies (makes true)
$\Vdash$	follows logically from
$\nVdash$	does not follow logically from
$\dashv \vdash$	is logically equivalent with
p	statistical probability
P	epistemic probability
$C(S)$ and $C_r(S)$	logical content and relevant logical content of a sentence set S
$E(S)$ and $E_r(S)$	empirical content and relevant empirical content of a sentence set S
Ω	possibility space
AL	algebra
$\ A\ $	semantic extension of formula A
$df(x)/dx$	first derivative of f(x)
$\sqrt{}$	square root
$\int$	integral
$\lim_{n \rightarrow \infty}$	limit for n going to infinity
$R(T)$	Ramsey sentence of theory T
$(IN) \text{ DEP } (X,Y)$	variables X and Y are probabilistically (in)dependent

Foreword

I have written *Philosophy of Science: A Unified Approach* for beginners as well as for graduate students and experts. This twofold orientation is reflected in the major structure of the chapters, which start with introductory sections for beginners, followed by sections entitled “complementary and advanced topics,” in which these topics are explained in a condensed way, including presentations of my own accounts. In this way, this book attempts to offer not only an introduction, but also a detailed overview of the various areas of the Philosophy of Science and their interrelations, which will help the reader familiarize him- or herself with the present state of the art in this field.

This book addresses itself both to philosophers and to scientists of all disciplines. By way of bridging Philosophy of Science with neighboring fields such as statistics, hermeneutics, cognitive science, the debate on science and values, and detailed examples of contemporary theories in sciences and humanities, this book tries to convey a deep understanding of the sciences and their methods. This book is *self-contained* and written without presuppositions about the readers. On rare occasions, formulas are used, but I’ve always explained them in a clear and easy-to-follow language.

Although this book introduces the reader to all major areas of philosophy of science, it also has a bias and a general philosophical message. Concerning the bias, the author understands himself as a follower of logically and empirically oriented philosophers of science such as Carnap, Hempel, Reichenbach, and Stegmüller in the German-speaking area. This book’s general message is reflected in its subtitle, *A Unified Approach*. Behind the diversity of phenomena investigated in disciplines as different as elementary particle physics and social sciences, or chemistry and linguistics, one can recognize a deep common ground that unites all of the sciences. This unity, however, is not so much a unity at the level of objects (as it was claimed by the proponents of physical reductionism), but rather a *methodological unity*. [Chapter 2](#) is concentrated on finding the common methodological features of the sciences and their epistemological presuppositions. The remaining chapters will not only work out these common features in all details but will also reveal some important differences between characteristic groups of disciplines, such as the physical sciences versus special sciences (e.g. biology, psychology, or social sciences).

Use in teaching: By restricting the teaching materials to Chapters 1–5 and omitting all sections on advanced topics, one obtains the basis for a semester-long two-hours-per-week course. Covering all of the material contained in this book will fill two such courses. The *core materials* of this book are highlighted in the form of *definitions* and *key propositions*.

Conventions on numberings: The numbering of figures follows the following pattern: “Fig. SecNo.-No.,” where “SecNo.” is the two-digit section number, and “No.” is the number of the figure within the section. Example: “2.1-2” is the second figure of [sec. 2.1](#). In the same way, *definitions* are numbered with “(Def. SecNo.-No.),” *key propositions* with “(Prop. SecNo.-No.),” and *examples* or other *highlighted* passages with “(SecNo.-No.).” Example: “(Def. 4.2-2)” is the second definition of [sec. 4.2](#); “([3.1-4](#))” is the fourth example of [sec. 3.1](#). *Quotations:* Double quotes are used for the purpose of literal quotation and single quotes for stylistic quotation (scare quotes).

At the end of every chapter, the reader finds hints on selected literature and exercises. This book concludes with a comprehensive bibliography and several indices.

This textbook is a revised and significantly extended English elaboration of my German Introduction to Philosophy of Science (“Einführung in die Wissenschaftstheorie”). The introduction has been developed from a series of lectures in the Philosophy of Science that I have given at the Universities of Salzburg, Erfurt, Düsseldorf and, less frequently, at other universities (including California at Irvine, Yale, Vienna, Helsinki, Graz, and Fribourg).

Parts of the text of the German edition have been pretranslated into English by Fergal Treanor. For valuable help concerning matters of content, I am indebted to Ioannis Votsis, Paul Thorn, Hannes Leitgeb, Wolfgang Spohn, Martin Carrier, Franz Huber, Markus Werning, Oliver Scholz, Andreas Hüttemann, Markus Schrenk, Hans Rott, Wolfgang Balzer, Gernot Kleiter, Theo Kuipers, Ilkka Niiniluoto, Erik Olsson, Stephan Hartmann, Jon Williamson, Igor Douven, Ken Gemes, James Ladyman, John Worrall, Stathis Psillos, James Hawthorne, Alan Musgrave, Mathias Frisch, Clark Glymour, Kevin Kelly, Horacio Arló-Costa, and Alvin Goldman. What remains on my side is to wish the readers of this book that it shall bring them much fun and intellectual profit.

Gerhard Schurz
Düsseldorf, March 2013

Foreword to the Second Edition

Since 2013, much has happened in the philosophy of science. Therefore, the Second Edition contains a lot of new content, concerning controversies about the role of values in science (sec. 2.10), results about meta-induction (sec. 2.11), debates concerning the question of scientific realism (sec. 5.8), and the mechanistic philosophy of science (sec. 6.4.1). In compensation, other and more technical parts have been shortened. Moreover, new literature has been added all over the place. The exercises with selected solutions have been extended. The index of subjects has been expanded so that it can be used as a *glossary* : for important terms, the page numbers on which the term is defined are shown in **bold**.

I would like to extend my thanks for valuable content-related help from the first edition to include Karel Lambert, Brian Skyrms, Branden Fitelson, Gila Sher, Johan von Benthem, Lorenzo Magnani, Holger Andreas, Mario Günther, Peter Brössel, Tom Sterkenburg, Leah Henderson, Richard Dawid, Matias Osta Vélez, Jochen Briesen, Gustavo Cevolani, Atocha Aliseda, Petr Spelda, Elke Brendel, Corina Strößner, Alexander Gebharder, Christian Feldbacher-Escamilla, Gottfried Vosgerau, Alexander Christian, Karim Baraghith, Niklas Parwez, Avril Styrman, and Christoph Schamberger.

Gerhard Schurz
Düsseldorf, September 2025

1 Introduction

Where Do We Stand?

DOI: [10.4324/9781003675075-1](https://doi.org/10.4324/9781003675075-1)

This chapter has been made available under a CC BY-SA license.

1.1 Tasks and Aims of Philosophy of Science

Science plays a determining role in the life of today's society. In what sense and to what degree can we trust the results of science? There is a scientific discipline, which examines such questions systematically: the *Philosophy of Science*, which in the German language is called "Wissenschaftstheorie" ("Theory of Science"). Following the German usage of the word, I always understand "science" ("Wissenschaft") in its *broad* meaning – including besides natural sciences, also social sciences, humanities, etc.

The discipline of philosophy (or theory) of science investigates how scientific knowledge works – its goals and its methods, its achievements and its limitations. Like any other scientific discipline, philosophy of science has historically developed out of philosophy. In the present days, it is practiced both by philosophers and by scientists of various disciplines.

One distinguishes between the *general* philosophy of science and philosophies of *particular* sciences. The latter ones are concerned with special kinds of disciplines such as philosophy of physics, biology, psychology, the social sciences, or the humanities. The general philosophy of science discovers those components of knowledge, which are more or less common to all scientific disciplines. Its main questions are:

- i. How is a scientific language constructed?
- ii. What are the rules of correct scientific reasoning or argument?
- iii. What are the distinctive features of a scientific observation or 'datum'?
- iv. What is a scientific law, and what does a scientific theory consist of?
- v. How are law hypotheses and theories empirically tested? When do they count as confirmed, and when as refuted?
- vi. What are the criteria for scientific progress?
- vii. What is achieved by a scientific explanation and what is causality?

This book focuses on these questions of the general philosophy of science, and the following chapters are arranged in corresponding order to the above questions. At the same time, we shall also deal with applications and topics of the philosophies of particular sciences, and examine characteristic *differences* between different branches of science. The most *general* questions of philosophy of science include:

- viii Is there an objectively discoverable reality and an objective truth?
- ix What connection exists between science and value judgments?

Question viii lies on the border between philosophy of science and epistemology (see [sec. 2.8](#)), and question ix on the border with meta-ethics (see [sec. 2.5](#), [2.10](#)).

Beyond its fundamental significance, which consists in providing answers to the above questions, philosophy of science also has a number of significant *applications*, both inside and outside science.

The *internal* applications of philosophy of science consist of the supply of foundational knowledge, which aids the scientist's decisions concerning new or controversial questions. Internal applications also consist of working out *interdisciplinary* common ground and providing *argumentative* competence in critical reflection. Not least, philosophy of science has often paved the way for new scientific disciplines (such as probability theory and statistics, measurement theory, or parts of cognitive science).

Among the *external* applications of philosophy of science to problems in a societal context, two in particular should be emphasized:

- a. The *demarcation* problem is highly significant in society. In this context, it consists of the question of which of our ideas have a claim to the status objective scientific knowledge that should be taught in public educational institutions, as opposed to subjective opinion, political values, ideologies, or religious convictions. This question became politically explosive in the controversy surrounding the creationism movement in the United States. For example, judge W. R. Overton in 1981 (McLean vs. Arkansas) and judge Jones in 2005 (Kitzmiller vs. Dover) based their decisions against the teaching of creationism in school on demarcation criteria between science and religious belief proposed by philosophers of science (cf. [Bird 1998](#), 2–9; [Branch 2011](#), 174). To tell the full truth, there were also philosophers of science, such as Laudan ([1983](#)), who have argued against the possibility of demarcating science from pseudoscience, and they have frequently been quoted by creationists as support for their position (to the embarrassment of other philosophers; cf. [Pennock 2011](#), 176ff).
- b. Equally significant is the critical role of philosophy of science concerning the commercial and political utilization of science, which counteracts the danger of *abusing* science and its results. Politicians, the media, and business leaders enjoy quoting expert opinions, which are presented all too often in a false or one-sided way in order to serve preordained purposes ([sec. 2.10.1](#), [4.3.6 – 7](#), [4.4](#)). But also applied scientists are often tempted to put subjective (e.g. financial) interests above objectivity. An example is the discovery of systematic biases in pharmaceutical research toward the financially supporting firms; philosophers of science have promoted scientific journal policies that may help to counteract such biases ([Brown 2008](#)).

1.2 On the Historical Development of Philosophy of Science

Although the term “philosophy of science” was only introduced in the 20th century, the philosophy of science is as old as science itself and had a parallel development (see [Losee 2001](#)). The history of the philosophy of science begins with *Aristotle* (384–322 B.C.), the ancient world's great systematizer of knowledge. Aristotle was much more oriented toward experience than was his teacher Plato. He was, nevertheless, like most philosophers after him, a follower of the *foundationalist* epistemological program. In this epistemological program, it is presumed that real knowledge is only possible if it stands on a foundation of certain and necessary principles, which are arrived at *not* through uncertain experience, but through rational intuition (cf. [Triplett 1990](#), sec. I; [Albert 1985](#), 18–28). Aristotle spoke in this context of “intuitive induction” ([Losee 2001](#), sec. Ia). In the current philosophy of science, by contrast, the *fallibilist* epistemological program has the upper hand, which concedes that our understanding of reality is basically *fallible*, and our scientific knowledge can be more or less well confirmed, but it cannot be guaranteed to be free of error. In a broader sense, of course, the scientific method is foundation-oriented, too, insofar it justifies its knowledge from most evident insights (primarily observation and logic) by inferences that are as reliable procedures possible – in contrast to coherentist epistemological programs, according to which the mutual coherence of beliefs is sufficient for their justification (cf. [Schurz 2024a](#), [sec. 2.1 – 2.2](#)).

Considerable advances in science and its philosophy had already been achieved by the late Middle Ages ([Crombie 1959](#)). Under the influence of philosophers such as *Roger Bacon* (1214–1292) or *William of Ockham* (1295–1349), the foundationalist program was gradually replaced by empirically oriented methodology. Of lasting importance is the principle of *Ockham's razor* , according to which one should reject all theoretical assumptions which are not needed to explain the observable. The decisive breakthrough, however, did not take place until the early modern period, the Philosophy of the Enlightenment. Building on the great successes of scientific method in this historical period – thanks to pioneers such as *Galileo Galilei* (1564–1642), *Isaac Newton* (1642–1727), and (later) *Charles Darwin* (1809–1882) – the influential *empiricist* movement was established, beginning with *Francis Bacon* (1561–1626) and *John Locke* (1632–1704) until *David Hume* (1711–1776) and *John Stuart Mill* (1806–1873). In the same period, the foundationalist epistemological program developed into the *rationalist* movement, prominently represented by *René Descartes* (1596–1650), *Gottfried W. Leibniz* (1646–1716), and *Immanuel Kant* (1724–1804).

The essential difference between the empiricist and rationalist philosophies can be characterized as follows. For the empiricists, sentences which may be justified *a priori* – that is, by pure reason and with rational certainty – are limited to so-called analytical sentences, that is, sentences whose truth is based on logic and conceptual convention. Such sentences possess no realistic content; they say nothing about the real world – only *synthetic* sentences may do this, but they can only be justified *a posteriori*, that is, based on experience. For the rationalists, on the other hand, *a priori* justifiable sentences with realistic content do exist, namely the so-called synthetic *a priori* sentences. However, both Descartes' and Kant's attempts to justify synthetic *a priori* principles of the sciences were unsuccessful and have been refuted by the further development of science. The development of the modern philosophy of science is closely bound with the insight that the foundationalist epistemological program was not tenable – at least not in its 'strong' version as explained above.

For a long time, classical systems of philosophical empiricism too had reductionist and foundationalist tendencies. The skeptical consequences of empiricism, which called for epistemic humility, were first recognized and radically formulated by Hume. Hume showed that the two central tenets of scientific method, the *causality* and the *induction principle* , are neither logically nor empirically justifiable. The causality principle says that all regular successions between events are produced by cause–effect relations; but Hume objected that all what one can observe are the successions of the events, while the idea that the earlier event or 'cause' produces the later event as its "effect" is metaphysical fiction (see [sec. 6.6.2](#)). The induction principle asserts that regular connections between events that have been observed in the past will also hold in the future; but Hume argued that this inference can neither be justified by logic nor by experience (see [sec. 2.11.1](#)). The problems involved in the justification of the principles of induction and causality have occupied philosophers up to the present day.

The apparent *insufficient justifiability* of empiricism put wind in the sails of subsequent rationalist movements, in which the foundationalist epistemological program, however, was gradually given up or replaced by pragmatized or historicized 'rationalisms' (see [sec. 1.3.2](#) –1.3.3). In the 20th century, *post-empiricist* and *post-rationalist* approaches have moved considerably closer to each other, and in the context of this tension, present-day philosophy of science has developed.

1.3 Philosophical Positions in Contemporary Philosophy of Science

1.3.1 Logical Empiricism

One of the most important origins of modern Philosophy of Science and Analytic Philosophy is logical empiricism, which was developed in particular by the Vienna Circle, but also by others, for example by members of the Berlin circle around Hans Reichenbach (1891–1953). The *Vienna Circle* was a group of philosophers and scientists in Vienna, whose core consisted of Moritz Schlick (1882–1936), Otto Neurath (1882–1945), and Rudolf Carnap (1891–1970). In the tradition of Ernst Mach (1838–1916), this group attempted to reestablish empiricism and scientific philosophy in general (cf. [Stadler 2001](#)). What was new about their situation was the development of *modern logic*, which had not been founded until the end of the 19th century. Modern logic made it possible to represent any language-based epistemological system with mathematical precision, and so the Vienna Circle hoped they had found at last the methodological equipment for a philosophy which could progress in a scientific manner (cf. [Schlick 1930/1931](#), 155f). In the middle of the 1930s, the Vienna Circle, with political leanings from liberal to socialist, had to disband under pressure from the National Socialists. The majority of its members emigrated, some via other countries, to the United States, where the logical-empirical movement, uniting with related Anglo-American directions, developed into Analytic Philosophy and gained a foothold in Europe in the postwar period.

Even today, many textbooks present logical empiricism as if it had represented a single dogmatic position; in reality, however, its proponents were highly capable of learning and continuously revised their positions. What today's philosophy of science can learn from logical empiricism is not so much any particular thesis but rather the high standards of conceptual and argumentative precision. In the phase up to 1935, the philosophy of science of logical empiricism was restricted by *positivism* and *reductionism*. In the later phase, the logical empiricists gradually rejected their restricted positions and replaced them with theses, to which the designations "empiricism" or "positivism" in the usual sense *no longer apply*. In one regard, logical empiricism has parted company with its classical predecessors early on: with the refutation of the classical empiricist thesis that at least 'direct' observation reports are infallible. In the so-called protocol sentence debate, Neurath and Carnap – but not Schlick – became convinced that even observation sentences such as "there is a table there" are in principle fallible. Empiricism, according to this new view, just meant that observation sentences have an epistemically favored role within the whole system of discovery (see [Carnap 1932/1933](#); [Neurath 1934](#), 113).

Empirical Reductionism, by contrast, was retained for quite some time. Classical empiricism advocated the following reductionist thesis: all 'serious' scientific concepts must be translatable via a chain of definitions into observational concepts. This program of reduction was adopted by logical empiricism in its early phase ([Carnap 1928a](#)). Later, however, the view prevailed that *theoretical concepts* such as "power," "electrical field," or "human character" *cannot* be defined by observational concepts, but rather go far beyond the immediately observable ([Hempel 1951](#); [Carnap 1956a](#)). In the so-called *standard* philosophy of science of the 1960s, the classical empiricist epistemological model had already broken down. What remained was a *minimal empiricism*, which consisted of the requirement that scientific theories must have empirical consequences, by means of which they can be tested.

A similarly modest empiricism is defended by contemporary empiricists such as van Fraassen (1980, [2002](#)). He does not only reject infallibilism and reductionism but also the *linguistic* orientation of logical empiricism. Instead, he defends a semantic conception of theories, in which theories are regarded not as sets of sentences, but as sets of conceptual models, similar as in structuralist philosophy of science (cf. [sec. 3.1.2](#)). Characteristic of van Fraassen's empiricism (1980, sec. 1) are two ingredients: first, a defense of the concept of theory-independent observation against Kuhn-followers (see [sec. 1.3.3](#)). And second, his rejection of scientific realism in favor of an instrumentalist position, according to which the theoretical (i.e. non-observable) part of a theory can only be reasonably said to be more or less *empirically adequate*, but not true or false. In this latter respect, van Fraassen is more restrictive than Carnap (e.g. 1928a) who always argued for ontological

pluralism , according to which both realism and instrumentalism are possible but epistemically undecidable practical attitudes toward scientific theories.

Two boundaries , which early logical empiricism had tried to maintain, became permeable in contemporary ‘minimal’ empiricism. *First* , the boundary between *science* and *metaphysics* became porous: it turned out that principles, which are of no empirical consequence *in isolation* , may give rise to new empirical consequences when joined with other theoretical sentences (see [sec. 5.4.2](#) ; [Hempel 1951](#) , 115; [Stegmüller 1969](#) ff, vol. II.1, 293–295). In this way, the *post-positivist* philosophy of science gained new access to the discussion of metaphysical principles such as realism and causality. Conversely, “inductive” branches of metaphysics emerged that sought to justify their principles not by pure thought but through empirical science (cf. [Hüttemann/Schurz 2025](#)). *Second* , the boundary between analytic and synthetic sentences became porous when it was recognized that the meaning of theoretical terms is not determined by isolated definitions, but by the respective background theory as a whole ([Carnap 1956a](#) , [1963](#)). In parallel fashion, Quine ([1951](#)) argued that in *natural* , empirically observed languages, no clear criterion exists by which one could undertake a demarcation between analytic and synthetic sentences.

As a consequence of this, analytic philosophy’s *linguistic* orientation (e.g. [Carnap 1950a](#)) was gradually given up. Linguistically oriented philosophers try to answer all philosophical questions by means of semantical concept analysis. But if a clear divide between (analytic) questions of meaning and (synthetic) questions of fact is missing, this exclusively linguistic orientation is no longer tenable. For this reason, Carnap ([1956b](#) , suppl. E) started to work on a system of pragmatics. Later on, Quine ([1974](#)) developed a *naturalistic* view, according to which science and philosophy form a continuum, the latter’s theses also having to be corroborated by experience.

1.3.2 Critical Rationalism

Also associated with the Vienna Circle was *Karl Popper* (1902–1994), the founder of so-called critical rationalism. From the beginning, Popper’s theory of science displayed the very non-reductionist tendencies that logical empiricism had accepted only after years in transition. According to Popper, scientific theories may go any distance beyond our experience, as long as they are empirically testable. Popper’s tests consist of *falsification attempts* , which should be as severe as possible. Popper supported his emphasis on falsification (instead of verification) by citing the *asymmetry* of *verification* and *falsification* of strictly universal hypotheses such as “all metals conduct electricity.” Such hypotheses cannot be verified by any finite number of observations, because they speak about a potentially infinite domain of instances – in our example, all metals in the past and in the future. However, these hypotheses can be falsified by a single observed exception ([Popper 1935](#) /2002, sec. I–IV).

For Popper, falsifiability was the decisive criterion of *demarcation* between science and speculation: scientific theories are not verifiable, but they must be falsifiable. Once falsified, they are *rejected* as false; but if they withstand manyfold falsification attempts, they are viewed as *corroborated* .

Popper’s falsificationism was also subjected to diverse criticism in the discussion which followed. Lakatos ([1970](#)) pointed out that scientific theories are almost never rejected on the basis of one single counterexample, but may be immunized against conflicts with recalcitrant data by means of modifications that are usually more or less *ad hoc* (see [sec. 5.6.1](#)). Popper’s *anti-inductionism* – that is, his thesis that science could progress without making any use of induction – also came in for harsh criticism (see [sec. 2.7.2](#)).

Popper’s critical rationalism is far removed from classical rationalism. Nowhere in Popper’s philosophy is it claimed that knowledge of reality can be grounded in non-empirical a priori intuition; on the contrary, Popper always rejected this. Thus, Popper’s philosophy is *post-rationalist* just as that of the late logical empiricists is *post-positivist* . Popper stressed more strongly than late logical

empiricism, however, that observation sentences are not only fallible, but also *theoretically biased* , so that the borderline between observational and theoretical terms cannot be sharply drawn (1935/2002, sec. V.30 addendum; new app. X). With this argumentation, critical rationalism opened its doors to the attacks from historical relativism, to which we turn now.

1.3.3 The Historical Account and Relativism

At the end of the 1950s, the historian of science Thomas Kuhn ([1962](#)) rattled the scientific world by presenting a fundamental attack on the standard philosophy of science. Factual science, said Kuhn, behaves differently from the normative ideals ‘prescribed’ by the philosophers of science. The alternative scientific model developed by Kuhn takes more of a historical-sociological than a logical-cognitive approach. According to Kuhn, scientific development takes place on the basis of so-called paradigms, such as the paradigm of classical (Newtonian) physics or (Darwin’s) theory of evolution. A Kuhnian paradigm contains at least the following three components: (i) very general theoretical principles or models, (ii) typical examples of successful applications, and (iii) methodological-normative assumptions (cf. [Kuhn 1977](#) ; [Hoyningen-Huene 1993](#) ; [Schurz 1998a](#) , [2014c](#)). The sociological correlate of a paradigm is a *scientific community* , a community of experts in the given field, who stick to a paradigm and work on its further development. According to Kuhn, a paradigm determines not only the fundamental principles, but also the major open problems and the *interpretation* of observed data – no, it even determines the observed data *themselves* , because, according to this view, *all observation is theory-laden* , and hence theoretically biased. For Kuhn, there is no observation which is independent of background theory or paradigm. Kuhn takes this strong thesis from Hanson ([1958](#)).

According to Kuhn, scientific development takes place in two successive phases, a *normal* and a *revolutionary* phase. Common acceptance of a paradigm makes the continuous growth of knowledge possible in the normal phase of a scientific discipline. When recalcitrant data, so-called anomalies, resist explanation by the given paradigm, the conflict is settled by more-or-less *ad hoc* modifications of current versions of theories within the accepted paradigm. Should such anomalies become more frequent, however, younger scholars will begin to look for a *new* paradigm. As soon as one is found, scientific development enters a revolutionary phase for a certain time, in which two paradigms struggle for supremacy. Kuhn’s examples in (1962) include the transition from Ptolemaic to Copernican astronomy and from Newtonian to Einsteinian physics. Kuhn claims that there are no common standards of rationality during a paradigm change and all empirical data are seen by the new paradigm in a new way. For this reason, the two competing paradigms are, according to Kuhn’s ‘notorious’ *incommensurability thesis* , rationally incomparable. So the struggle for supremacy does not take place in the form of a comparison of cognitive achievement, but rather in the form of a *power struggle* , in which supporters of the old paradigm eventually die out, as a result of which the new paradigm prevails, and a new normal scientific phase is heralded in.

The standard philosophy of science was unable to effectively counter the argument of the theory-dependence of scientific observations and its relativistic consequences. For all questions about the cognitive nature of observation had been banned by the narrowed perspective of the *logic* of science. For Popper, the question of what should be called an observation sentence was a question of pure convention (1935/2002, sec. V.30). The logical empiricists saw it in much the same way. It will be shown in [sec. 2.8.2](#) that a different picture results as soon as we take on an *empirical-cognitive* perspective. After all, if there were no difference between what experience *shows* , and what theory leads us to *expect* , then empirical science would be an exercise in *permanent self-deception* . Any theory of the empirical sciences which cannot explain this central difference must be extremely unsatisfying. Moreover, the loss of a viable criterion of empiricism would also undermine any credible distinction between real and fake facts. After right-wing movements began to criticize scientifically established facts, for example, concerning the climate crisis, as “leftist ideology,”

philosophers with left-leaning views tend to embrace once again the experience-based objectivity that was criticized by them some decades ago as an expression of “positivist ideology.”

Kuhn’s alternative model of science was the subject of much controversial discussion in the decades that followed. As well as an opposition to Kuhn, both a moderate and a more radical branch of Kuhn-followers was formed (cf. [Hoyningen-Huene 1993](#) , 212f). The more radical group gave relativism a strong impetus – most prominently in Feyerabend’s “Anarchist Epistemology” (1975), who with his maxim “anything goes” argued against methodological rules that are guiding or should guide science. The moderate Kuhnians, by contrast, aimed to separate Kuhn’s insights in the philosophy of science from the exaggerations of his “radical interpretation.” Kuhn himself later softened his incommensurability thesis and granted limited recognition to the possibility of rational comparisons between paradigms. ¹A moderate Kuhn-follower, and a Popper-follower at the same time, is Imre Lakatos and his conception of scientific research programs. Another moderate Kuhn-follower is Larry Laudan ([1981](#)), who argued that the empirical success of a scientific theory is not necessarily a reliable indicator of its truth. Such moderately Kuhnian developments have brought important insights to light for the philosophy of science (see [sec. 5.6](#) and 5.8). A further controversy centered around Kuhn’s thesis that in its normal scientific phases a scientific discipline is always dominated by a single paradigm. Scientists and philosophers of science have pointed out that, contrary to Kuhn’s thesis, disciplines outside of physics often exhibit a persistent coexistence of rival paradigms (cf. [Schurz/Weingartner 1998](#) ; [Kornmesser/Schurz 2014](#)).

1.3.4 Metaphysical and Scientific Realism

In reaction to the Kuhnian challenges of the logical-empirical model of science, many philosophers have suggested returning to the metaphysical approaches of realism and necessity, from which the philosophy of science in the early 20th century had tried to free itself. For example, it is a consequence of post-positivist philosophy of science that the term “mass” actually has different *meanings* in Newtonian and Einsteinian physics. This theory-dependence was also one of the main pillars of Kuhn’s incommensurability thesis. The early Putnam (e.g. 1975b, 219ff, 229–235, 240f), together with Kripke ([1972](#) , 55–59), developed the following argumentation against this consequence: scientific progress in the sense of a movement toward objective truth is only possible if the *reference* of scientific terms, that is, their relationship to reality, *is rigidly fixed* with metaphysical necessity, in the sense that this reference stays the same in all possible theories and across all possible worlds. Kripke and Putnam assumed the reference fixation to take place by acts of baptizing, together with a rigid sameness-relation between natural kinds, with the result that statements such as “water is H_2O ” become true by “metaphysical necessity” (see also [Shoemaker 1980](#)). Fine ([1994](#)) also regards such identity relations as metaphysical necessities constituted by essential properties.

We agree with Kripke and Putnam that the meaning of natural kind terms such as “water” cannot be learned by an empirical definition (as in [Carnap’s 1956b](#) account), but by acts of ostension (“this is water”) grasping the right sameness-in-kind relation between water exemplars. But as a number of arguments have shown, scientific knowledge cannot determine sameness-in-kind relations and the reference of (theoretical) concepts with metaphysical necessity. ²Of course, classification systems for scientific (e.g. chemical or biological) kinds can be fixed *analytically* – that is, through definitional conventions (cf. [Sidelle 1989](#)). However, which classification system accurately reflects reality depends on the theoretical background knowledge ³which undergoes historical change ([Hendry 2008](#)). The idea that a necessary relationship between language and experience can be established independently of experience is in the end a regression toward a foundationalist epistemic model. In fact, the later Putnam ([1977](#) , [1982](#)) relinquished his earlier metaphysical realism by arguing as follows: if neither our empirical nor our theoretical knowledge can determine the reference of our concepts in a unique way, then *nothing* can do this – the consequence is an

unavoidable “indeterminacy of reference.” The same indeterminacy of reference has been worked out by Quine, but while Quine remained a naturalistic realist, the late Putnam adopted a more-or-less idealistic position that he called “internal realism.”

An account of necessity that is metaphysical in a weaker sense than the above ones is that of Armstrong ([1983](#)). According to him, laws of nature supervene on an assumed necessitation relation between “universals” (such as “water” and “H₂O”), which is, however, *contingent* – that is, these necessities may be different in different possible worlds ([sec. 6.5.6](#)). A similar middle ground between metaphysics and empiricism is Lewis ([1973](#) , [1986a](#)), who developed a possible worlds semantics for counterfactual and causal conditionals, but at the same time defends an elaborated Humean account to laws of nature, according to which laws supervene on the ‘mosaic’ of empirical facts ([sec. 6.6.5](#) ; cf. [Schrenk 2017](#) , [sec. 4.3](#) , [5.3](#)).

Like metaphysical realists, proponents of scientific realism also advocate for the convergence of science toward truth – but not on the basis of metaphysical postulates, but based on the empirical success of science. To support this position, scientific realists require robust arguments that allow them to infer the approximate truth of a theory from its empirical success. The most well-known argument of this kind is Putnam’s *no miracles* argument (1975a, 73). According to this argument, the empirical success of a theory would be a sheer miracle unless one assumes that the theory is at least approximately true. The no miracles argument has been subject to extensive criticism, notably by Laudan ([1981](#)), who, in his *pessimistic meta-induction* , argued that virtually all empirically successful theories of the past have eventually been shown to be false, whence it is inductively highly probable that our current best theories will suffer the same fate. Other philosophers of science, such as Boyd ([1984](#)) and Psillos ([1999](#)), have defended scientific realism against such anti-realist arguments, leading to a scientific realism debate that continues to this day (see [sec. 5.8.3](#) – [5.8.8](#)).

1.3.5 Naturalism and the Cognitive Turn

Modern naturalism goes back primarily to Quine ([1951](#) , [1960](#)), who rejected the analytic-synthetic distinction, thereby dissolving the clear boundary between scientific philosophy and empirical science. For Quine, naturalized epistemology becomes an empirical discipline, which is “scientifically investigating man’s acquisition of science” ([Quine 1974](#) , 3). The naturalistic program was emphatically taken up by a number of young epistemologists and philosophers of science (e.g. [Kitcher 1992](#) ; [Papineau 1993](#) ; [Kornblith 1994](#)).

Philosophical naturalism found expression in several philosophical programs. A first concretization of naturalism is *evolutionary epistemology* (e.g. [Campbell 1984](#)): many aspects of humans’ cognitive capacities and achievements can be explained by facts of evolution, in particular regarding perceptual knowledge. It has to be critically remarked, however, that naturalistic approaches themselves make certain epistemological assumptions which are *in need of justification* , for example concerning realism and induction, and so they cannot entirely replace classical epistemology. There is, however, an epistemological school of thought that indeed amounts to a *renunciation* of justification in this regard, namely epistemological externalism. According to its most well-known version ([Goldman 1986](#)), a belief counts as justified if it is produced by an “objectively reliable” method, without requiring anyone to be able to justify the reliability of the method in question. A justification of inductive prediction methods is not sought; instead, it is simply asserted that a predictive method is justified if it is reliable, that is, if its predictions will come out as true. Otherwise, it is considered unreliable. The externalist treatment of realism works in a similar way. However, since all nonscientific ‘methods’ – ranging from creationist doctrines of salvation to clairvoyance – also claim reliability for themselves, externalist epistemology fails to demarcate science from nonscience and thereby condemns itself to epistemic uselessness (see [Schurz 2024a](#) , [sec. 2.4](#)).

A second concretization of naturalism is the so-called cognitive turn. This turn resulted from the fact that meanwhile the cognitive sciences had attained the status of an interdisciplinary subject area in their own right, in which the basis of humans' cognitive and scientific capabilities is examined not only logically in philosophy but also empirically in psychology or linguistics and computationally in Artificial Intelligence research (cf. [Thagard 1996](#) ; [Anderson 2000](#)). As a result, several philosophers have started efforts to bring the approaches of the cognitive sciences into philosophy of science and vice versa (e.g. [Goldman 1986](#) ; [Bechtel 1988](#) ; [Bishop/Trout 2005](#)). The recent movement of "experimental philosophy" has even started to examine philosophical intuitions by empirical investigations instead of the traditional "armchair method," with interesting results concerning cultural differences ([Machery et al. 2004](#) ; [Knobe/Shau 2008](#) ; [Horvath/Koch 2020](#)).

The third specification of naturalism is, finally, naturalistic philosophy of science. Like Quine, it assumes a continuity between the methods of the empirical sciences and those of the philosophy of science. Questions of value are not excluded from science, but they are understood in an instrumental way – as questions of the optimal choice of means for given ends ([Laudan 1987](#)) – which closely aligns with the value-neutrality approach discussed in [sec. 2.5](#) . By contrast, the influence of externalism must be viewed critically: the philosophical aim of justifying fundamental scientific methods in contrast to pseudoscience is largely abandoned, and the reliability of scientific methods and results is simply presupposed ([Godfrey-Smith 2003](#) , 151). This runs counter to the aim of theoretical clarification of science and the exposure of the ideological misuse of scientific authority, as outlined in [sec. 1.1](#) .

1.3.6 Structuralist and Semantic View of Theories

In search of a better defense against the Kuhnian offensive, Stegmüller ([1976](#) ; [1977](#) , vol. II., sec. 6; 1979b) suggested replacing the "statement view" of the standard philosophy of science with a so-called non-statement view. Stegmüller's suggestion is based on the idea, traceable to Suppes ([1957](#) , 246ff) and Sneed ([1971](#)), that it is better to reconstruct scientific theories as systems of set-theoretical *models* than as systems of statements in a logical language. Such systems of models are also called *structures* : hence, the name "structuralist" philosophy of science. In 1976, [Stegmüller \(1976](#) , 105) envisions this step as a radical upheaval. Later, he is more modest and argues for the transition from logical to set-theoretical formalization methods for "pragmatic and psychological reasons" concerning the simplicity of formal reconstructions (1979b, 65). For set-theoretical models are also characterized by statements, namely by statements in the language of set theory. Therefore, there exist straightforward possibilities of *translation* between the theory representations in the statement view and the non-statement view (see [sec. 3.1.2](#) ; [Schurz 2014b](#)).

While the statement-versus-non-statement debate turned out to be a bit of a red herring, the structuralist philosophy of science can claim significant achievements in the *detailed reconstruction* of individual scientific theories and theory nets (cf. [Balzer et al. 1987](#) ; [Balzer/Moulines 1996](#)). Moreover, the structuralist philosophy of science has received support from Anglo-Saxon philosophers such as van Fraassen (1980, sec. 2; 1989, sec. 9) and Giere ([1988](#) , [2004](#)), who called this position the "semantic view of theories" and developed it in a less formal way (cf. [Suppe 2000](#)). The semantic view of theories essentially consists in the claim that theories should be understood as models – a view later taken up by the "new mechanists" in the philosophy of science (e.g. [Glennan 2017](#) ; [Glennan et al. 2022](#) ; see [sec. 6.4.1](#)). However, the concept of a model remains vague: virtually anything can serve as a 'model' – from images to mathematical equations – as long as it serves to represent some aspect of reality (see [Glennan 2017](#) , 60; for criticism, see Chakravartty 2001; [Schurz 2014b](#)).

1.3.7 Pragmatic Philosophy of Science

In the 1980s, several authors spoke of a *pragmatic turn* in philosophy of science ([Stegmüller 1969](#) ff, vol. I, sec. XI; [Böhler et al. 1986](#) ; [Bernstein 2001](#)). Unfortunately, however, pragmatic philosophy and the use of the term “pragmatics,” in general, fall into two *opposing* camps. While one camp (e.g. Ch. S. [Peirce 1931](#) ff, van Fraassen 1980; [Rescher 1998](#) , [2001](#) ; [Schurz 1998](#) b) understands pragmatics in the *epistemically internal* sense, the other camp (e.g. [James 1907](#) ; [Rorty 1982](#) ; [Stich 1990](#) ; partly [Putnam 1995](#) ; Bernstein 2010) sees pragmatics in an *epistemically external* sense.

For the epistemically external camp, the pragmatic components of discovery are essential, but have nothing to do with its truth. By contrast, for the epistemically internal camp, it is these very pragmatic components, which are supposed to establish an objective conception of truth in a noncircular sense. The epistemically external camp makes knowledge (or justified belief) dependent on *any* kind of interests, be it predictive accuracy, political control, or emotional well-being, and sees therein a compelling argument for the unavoidable subjectivity of knowledge. In this sense, Stich ([1990](#) , 131ff) has argued that there are so many different conceptions of truth as there are different cognition-guiding interests. By contrast, for the internal camp, the reference to epistemically *internal* purposes is the very thing which makes possible a noncircular foundation of knowledge. For example, the reference to the cognitive purpose of maximizing true predictions enables a justification of induction that escapes Hume’s skeptical challenge ([sec. 2.11.2](#)). In conclusion, the external and the internal camp *differ so greatly* in their understanding of pragmatics, that they should not be subsumed under a single term.

1.3.8 Hermeneutics and Critical Theory

The *Philosophy of the Humanities* , which originated in the 19th century, goes back to Friedrich Schleiermacher (1768–1834), Wilhelm Dilthey (1833–1911), and Wilhelm Windelband (1848–1915). It started as a movement for the methodological demarcation of the humanities against the expanding natural sciences. Central to this philosophy is the use of *hermeneutics* , the study of interpersonal *understanding* (*Verstehen*) , for the purpose of such a methodological demarcation. The *methodological dualism* of the philosophers of the humanities has been brought to a head with the following thesis: in the natural sciences, we *explain* , but in the humanities, we *understand* . While scientific explanations are based on general law hypotheses, it was claimed that understanding in the humanities needs no recourse to such law hypotheses, because the human mind is not subject to strict law mechanisms, and understanding has direct access to the human mind.

Hermeneutics is part of a tradition from well before Schleiermacher, in which there was no mention of any rigid separation of the humanities from logic and natural sciences (see [Scholz 2001](#) , part I.B). The methodological dualism originated in the 19th century and had a strong influence on methodology debates of the 20th century (cf. [Von Wright 1971](#) ; [Manninen/Tuomela 1976](#) ; [Apel 1984](#)). The contemporary controversy in the humanities and social sciences about *quantitative* versus *qualitative* methods is partially a revivification of the old methodological dispute (cf. [Lamnek 1988](#) ; [Denzin/Lincoln 2000](#)). In response, the analytic philosophy of science has produced a number of approaches, in which the unity of hermeneutics and empirical methods is worked out (see [sec. 6.4.2](#) ; cf. [Hirsch 1973](#) ; [Follesdal 1979](#) ; [Taylor 1980](#) ; [Bühler 2003](#)).

Critical theory was developed by the philosophers of the so-called Frankfurt school, mainly Horkheimer, Adorno, Marcuse, and Habermas. It combines the methodologically dualist version of hermeneutics with the doctrine of the unavoidable interest-dependence of discovery and knowledge (see [Habermas 1966](#) ; [Apel 1984](#)). Critical theory also influenced the philosophy of science debate on the value-freedom versus value-ladenness of scientific research. We turn to this question in [sec. 2.5](#) and 2.10, where a modernized version of scientific value neutrality is developed – one that is compatible with value commitment in science and capable of withstanding relevant objections.

1.3.9 Radical Constructivism, Postmodernism, and the Science Wars

Radical constructivism was developed in the 20th century by biologists such as Maturana and Varela ([1992](#)) and psychologists such as von Glasersfeld ([1995](#)); philosophically, however, it is much older. The central argument of constructivism infers that, because our perceptions and descriptions of reality are constructed, there is no subject-independent reality – instead, *reality itself* is said to be a human construction. Even the early Kuhn embraced a constructivist position, claiming that “the world itself” has changed following a scientific paradigm shift ([Kuhn 1962](#) , chapter 10; see [Godfrey-Smith 2003](#) , 96).

In [sec. 2.8.2](#) , we will show that the argumentation of radical constructivism is based on a *fallacy* . Regardless of this, radical constructivism, in combination with the hermeneutic methodological dualism, has exerted its influence in the cultural and social sciences. There, it shaped French poststructuralism and the philosophy of postmodernism ([Lyotard 1984](#)), which reject scientific reality and its impact on human culture, instead promoting the idea that reality is entirely socially constructed. In the United States, these currents gave rise to the field of “Science Studies” and encouraged a left-leaning politicized understanding of science, as exemplified by feminist philosophy of science ([Harding 1986](#) ; cf. [Godfrey-Smith 2003](#) , sec. 9.1–9.4). Most pointedly, the social constructivist position can be found in the “science studies” of Latour and Woolgar ([1986](#)) and Knorr-Cetina ([1981](#)), where scientific laboratories (e.g. in nuclear physics or biochemistry) are studied by ethnographical methods of studying “foreign cultures” without engaging with the actual epistemic content being investigated, in order to learn something about the “social construction of scientific reality.”

The disregard of methodological standards and the misuse of scientific terminology met with strong opposition from the scientific community and led, in 1996, to a “counterattack” in the form of the *Sokal Affair* . In order to expose the lack of qualifications behind the wordplay of postmodern authors, Sokal ([1996](#)) published a hoax article titled “Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity,” which consisted of a jumble of quotes from postmodern thinkers. Despite its nonsensical content, the article was accepted and published by the postmodern-oriented American journal *Social Text* . When Sokal later revealed in another journal that his article had been a parody of postmodern discourse, the scandal was complete (cf. [Godfrey-Smith 2003](#) , 146f). The subsequently published book by Sokal and Bricmont ([1998](#)), entitled *Fashionable Nonsense: How the Postmodernists Abuse Science* , caused a major stir and became a catalyst for the so-called Science Wars ([Ross 1996](#)).

1.4 Introductory Literature and Exercises

1.4.1 Introductory Literature

A historically oriented introduction into the philosophy of science of the 20th and 21st centuries is Godfrey-Smith ([2003](#)). Ladyman’s introduction (2002) is focused on natural science and Rosenberg ([2008](#)) introduces to the philosophy of social science. Okasha ([2016](#)) gives a systematic but very brief introduction. An excellent introduction to the history of the philosophy of science is offered by Losee ([2001](#)). Stadler ([2001](#)) presents an extensive study of the Vienna Circle and its logical empiricism. For Popper’s philosophy of science, see [Keuth \(1998](#) , ed.). Hoyningen-Huene ([1993](#)) offers an introduction into the philosophy of Kuhn. [Hüttemann \(2021\)](#) offers an introduction into the metaphysics of science; for critical discussion of metaphysical accounts, see Ladyman and Ross ([2007](#)). On evolutionary epistemology, see Campbell ([1984](#)). The history and current situation of hermeneutics are dealt with by Bruns ([1992](#)).

1.4.2 Questions and Exercises

Sec. 1.1–1.2

1. What are the seven main questions in the general philosophy of science?
2. What is the major difference between the empiricist and rationalist ways of thinking in epistemology?

Sec. 1.3

1. What is understood by empirical reductionism?
2. What criterion of demarcation is suggested by critical rationalism?
3. What is a paradigm for Kuhn? Discuss Kuhn's incommensurability thesis.
4. How does scientific instrumentalism differ from scientific realism?
5. Discuss the methodological dualism of the philosophers of the humanities from a contemporary viewpoint.

Notes

1. Kuhn ([2000](#) , 33–57); cf. Okasha ([2016](#) , 87); [Seselja/Straßer \(2013](#) , 14f). Empirical data show that history of science proceeds much more cumulative than Kuhn suggests (see [Krauss 2025](#)). ↵
2. See [sec. 5.8.2](#) . Cf. Mellor ([1980](#) , 118f); Bird ([1998](#) , 108–120); Ladyman/Ross ([2007](#) , 290ff). ↵
3. In philosophy of science, “background knowledge” is not understood in the strict epistemological sense: it contains scientific beliefs that are rationally accepted but not necessarily true. ↵

2 The Question of Unity

In Search for Common Foundations of the Sciences

DOI: [10.4324/9781003675075-2](https://doi.org/10.4324/9781003675075-2)

This chapter has been made available under a CC BY-SA license.

2.1 Normative or Descriptive? The Method of Rational Reconstruction

Prima facie, we can identify two opposing views of the goals and methods of the philosophy of science:

1. According to the *normative* view, it is the goal of philosophy of science to say what science *should* be, and how it should be pursued. For this purpose, philosophy of science must determine in what scientific rationality consists, and on the basis of which criteria a scientific hypothesis may be justified.
2. According to the *descriptive* view, on the other hand, it is the task of philosophy of science to say what science de facto *is*, and how it proceeds. For this purpose, philosophy of science must describe and explain the existing sciences in terms of their historical development and their current structure.

The normative view is historically older, and currently still enjoys a wide following. Both logical empiricists and critical rationalists held this view. It was not until Kuhn ([1962](#)) set off the historical turn in philosophy of science that the descriptive counterpoint became popular. It can be found in the work of historically oriented philosophers of science (cf. [McMullin 1970](#) ; [Laudan 1977](#) , 162), of the structuralists (e.g. [Stegmüller 1976](#) , 267ff; [Balzer et al. 1987](#) , xvff), and of the naturalists (cf. [Giere 1999](#) , 160–163; [Ladyman 2002](#) , 4). So which position is right?

The position of the normativists can be traced back to Hans Reichenbach's ([1938](#) , 6f) distinction between the *context of discovery* , that is, the factual *genesis* of scientific knowledge, and its *context of justification* . The plausible prima facie argumentation of the normativists may be broken down as follows (e.g. [Popper 1935](#) /2002, [sec. 1.2](#)):

(2.1-1) *Normativist argumentation* :

Premise 1 : To assess the epistemic value of a hypothesis, *only* the question of rational justifiability of its truth claim is relevant.

Premise 2 : Questions about the hypothesis's historical development, the causes for its discovery, etc., are irrelevant to this purpose.

Conclusion : Therefore, philosophy of science should only examine science's rational justification and not its context of genesis – the latter being the remit of the history, sociology, or psychology of science, but having nothing to do with its philosophy.

Since Kuhn ([1962](#)), the argumentation of the normativists has been criticized in a very *sweeping* way. Here, we will attempt to clarify what is correct and what is flawed in this line of argument. The argument is certainly logically correct – assuming we take it to be analytically true that the philosophy of science is concerned with the assessment of the epistemic value of hypotheses. That is to say, *if* the premises of the argument were really true, then the conclusion would also be true. There is also *no error* in premise 1 – this thesis also seems to be analytically true, to the extent that science is concerned with the discovery of truth, and nothing else. But premise 2 is already infected by the 'germ' of the normativist fallacy. Because the

historical genesis of a scientific hypothesis includes not only so-called external factors which are epistemically irrelevant, but also internal factors which are epistemically relevant: for example, the experimental methods, which scientists de facto use, and the rational hypothesis evaluation, which they de facto apply.

The organic chemist Kekulé, for example, reported that his groundbreaking hypothesis of ring-shaped benzene molecules first appeared to him in a *dream* (see [Hempel 1966](#) , 16). This fact was, of course, epistemically irrelevant and played no role in the justification of the hypothesis. By contrast, the *Boyle-Mariotte law of gases* was discovered by means of the systematic extrapolation of data ([Langley et al. 1987](#) , 81) – and this fact played a decisive role in the justification of the hypothesis.

Philosophers should *learn* from the de facto methods and arguments of science as far as they are successful, and they can only do this by observing them. Instead of premise 2, all the normativists usually want to say, or at least all they should say, is the following *almost trivial* weakening:

(2.1–2) *Premise 2** : The epistemically *external* part of the history of the development of a hypothesis is irrelevant to the assessment of its epistemic value.

Of course, the conclusion of argument ([2.1-1](#)) no longer follows if premise 2 is replaced by the weaker premise 2*. And with this conclusion, the error of the normativist argumentation becomes *grave* . To realize this, we must simply ask the following question: how can philosophers of science *acquire* the appropriate rules and criteria for the scientific method? It would be naïve to presume that these rules and criteria might be gained through pure logic or mere intuition; science is much too complex for that. That would be like somebody trying to write a bridge-building manual without having examined the de facto practice of bridge building and typical examples of existing bridges. The philosophy of science, then, *can* and *should* learn from the real examples provided by the history of science. This brings us to the following result: although it is the *primary* task of the philosophy of science to explain the methods and criteria of *rational* science, it can only fulfill this task by *also* studying the *factual sciences* in their historical development.

Kuhn's criticism crystallized into the insight that the scientific models provided by logical empiricism and critical rationalism were too *simplicistic* to give a faithful picture of real science (cf. [Stegmüller 1979b](#) , 5; 1969ff, vol. II.3, 20). Reacting to this, young theorists proclaimed that the philosophy of science should even be restricted to the descriptive analysis of the sciences. But this position too is overstated. The question of the definition and criteria of scientific rationality must, of course, remain central to philosophy of science, as it is this question which defines philosophy of science as a discipline and sets it apart from the history and sociology of science. In summary, the philosophy of science is a discipline, which contains both descriptive and normative elements – an insight that has been worked out by [Stegmüller \(1969](#) ff, vol. IV.1, 9ff).

The philosophy of science's method may best be described as a *rational reconstruction* . My understanding of this method is different from Lakatos' understanding of it, which has been sharply criticized by Laudan ([1977](#) , 167–170). For Lakatos ([1971](#) , 119f), 'history' as understood by rational reconstruction is not the *actual* history, but a projection of ideals of rationality into history, and a selection of those parts of history that conform to that ideal. By contrast, in my understanding, the actual history forms one pole of the method.

More precisely, rational reconstruction moves between two poles: a (so-called) *descriptive corrective* , which should be adequately reconstructed, and a *normative corrective* , which contains rational norms. The descriptive corrective does not contain every piece of factual sciences. It contains only typical examples of de facto successful (i.e. unanimously and sustainably accepted) scientific discoveries. Moreover, it also contains typical 'counterexamples,' that is, examples of de facto refuted scientific hypotheses. Philosophy of science bases its reconstructions on these consensually accepted examples of scientific success or failure, though only in a preliminary and defeasible way.

The normative corrective consists of certain epistemic intuitions that are normally assumed by scientists, in spite of their various cognitive differences. That such a common ground exists is the position that will be defended in this chapter. First of all, the normative corrective contains one supreme epistemic goal or norm, which may be explicated as follows (cf. [Weingartner 1978](#) , [sec. 3.2](#)):

(Prop. 2.1-1) The *supreme epistemic goal* (G) of science is to find *true* and *content-rich* statements, laws, or theories, relating to a given domain of phenomena.

A sentence is the richer in content, the more *consequences* it has (see [sec. 3.7](#)). The restriction of the goal of science to true *and* content-rich statements is significant, because the probability and the content of hypotheses are often *inversely proportional*. We can maximize the truth chances of hypotheses by sticking to trivial tautologies like: “The Sun revolves around the Earth, unless it doesn’t.” Inversely, it is easy to present very substantial and impressive hypotheses if we disregard their truth value, for example, “I have invented a perpetual motion machine.” The true art of the scientist is to formulate hypotheses which have both empirical truth and substantial content and consequences.

In the above explication of the supreme goal of science, there appears the little word “truth.” One might object that the goal of truth is not metaphysically neutral. However, we understand this concept in a liberal way, leaving room for different metaphysical interpretations (see [sec. 2.8.1](#)). On the other hand, we don’t want to be too liberal about the meaning of “truth.” For people have rather different ideas of what truth is. While one person seeks truth in the platonic heaven, another person may equate truth with empirical success, and for a third one truth is what makes herself happy. To ensure that the goal of scientific discovery as explicated in [prop. 2.1-1](#) is not entirely vague concerning the meaning of “truth,” we need a *minimal* explication of what “search for content-rich truths” is intended to mean. This is provided by the second ‘normative’ component that is normally assumed by scientists as well as by philosophers of science, namely a minimal common epistemological model of scientific discovery. In [sec. 2.2](#), we will explain this minimal epistemological model of the sciences in detail. The key points are given here: it contains (1) a *minimal realism*, (2) *fallibilism* and critical attitude, (3) the struggle for *objectivity*, (4) a *minimal empiricism*, and (5) the struggle for *logical clarity*. Although this minimal epistemological model consists primarily not of normative, but of *descriptive* claims, we still regard it as part of the normative corrective of philosophy of science, as it gives clear meaning to the supreme goal of science, and is usually seen as a precondition of science.

Given the above described normative and descriptive correctives, the method of rational reconstruction may be precisely characterized as follows:

(Def. 2.1-1) *Method of rational reconstruction* : Develop generalized and logically precise *models* of scientific discovery, which on the one hand *fit the descriptive corrective*, that is, provide optimal explanations for the success of the examples and the failure of the counterexamples from the history of science, and which on the other hand may be *justified according to the normative corrective*, as providing optimal *means* of reaching the supreme goal of scientific discovery, given the assumptions of the minimal epistemological model.

In short, the philosophy of science abstracts generalized models from the descriptive corrective, which at the same time are (or should be) in line with the normative corrective. This is summarized in the schematic diagram in [fig. 2.1-1](#).

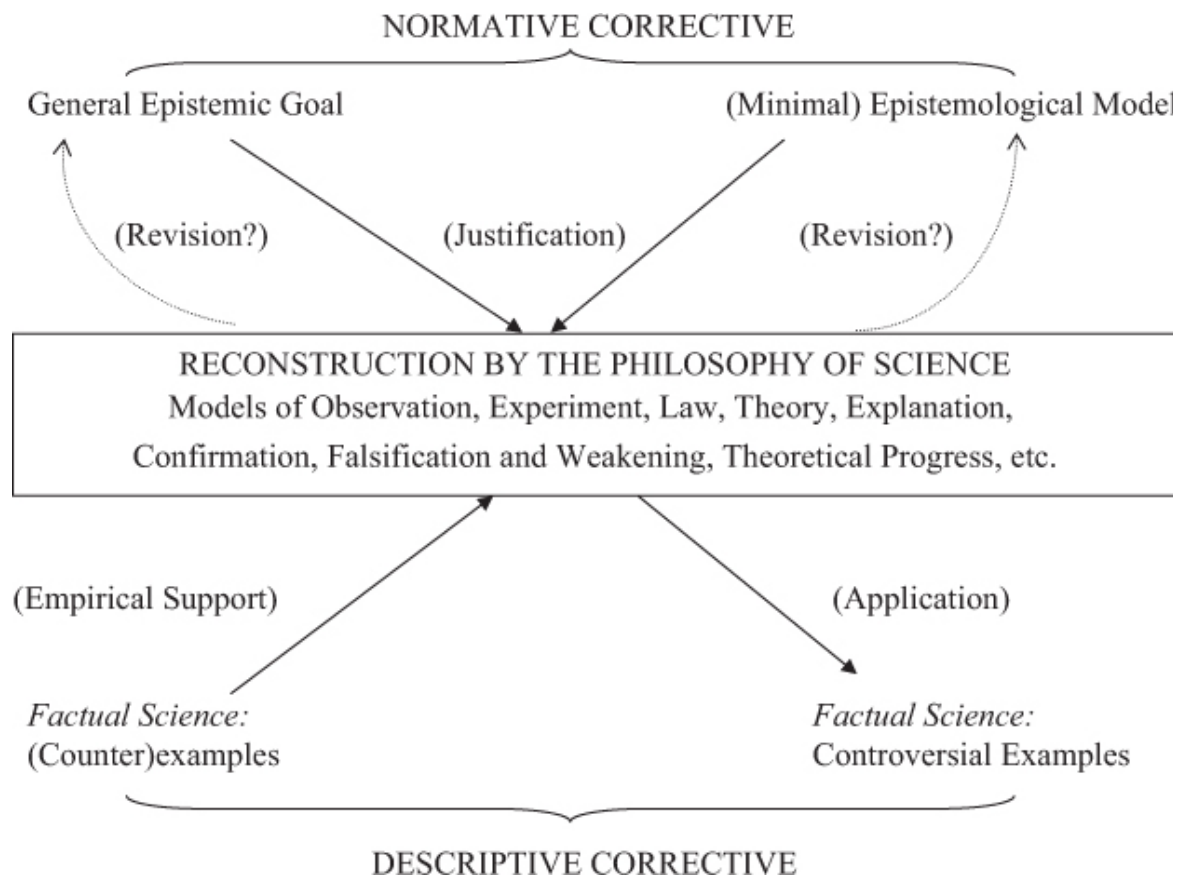


Figure 2.1-1 The method of rational reconstruction. [↗](#)

As long as both aspects of rational reconstruction – descriptive adequacy and normative justifiability – can be made to work in harmony with each other, the work of scientific reconstruction can be counted as successful and its results as *corroborated*. Of course, we cannot expect that everything we find in *factual science* will also be *rational* according to the normative corrective. For this reason, the descriptive corrective of the philosophy of science contains only *accepted* (counter) examples. When models developed by the philosophy of science have been corroborated, we can take the next step of applying them to *controversial* areas of scientific disciplines, in which no consensus has been reached – be it the objectivity problem of quantum mechanics, the teleology problem in biology, or the explanation-understanding problem in the humanities. It is exactly here that philosophy of science can supply valuable decision-making aids to the individual sciences. So, on the one hand, the philosophy of science learns from the individual sciences and is corrected by them; on the other hand, its reconstruction work can be instructive to the individual sciences and thus help them solve specific problems.

The goal of rational reconstruction, though, bringing together descriptive analysis and normative justification, does not always have to succeed. This applies, for example, when philosophers of science proceed from unrealistically high or one-sided normative standards, which will persistently disallow them to reconcile their normative corrective with the descriptive analysis of science. In such cases, the philosophy of science must go on to *revise* its own normative corrective. When it does this – this has already happened several times in the history of the philosophy of science – then the philosophy of science, to borrow Kuhn's terminology, is no longer in a normal scientific phase, but in a revolutionary phase.

To conclude this section, we will compare the model of rational reconstruction with two similar characterizations of the method of philosophy of science. Carnap (1950b, sec. 2) described the procedure of the philosophy of science as *explication*. Here, a vague concept of everyday or scientific language (the explicandum), for example, the idea of probability, is replaced by a precisely defined term (the explicatum), whereby the explicatum must differ from the explicandum insofar as it eliminates vagueness and ambiguity,

but it should also remain as close to it as it can and be as fruitful and simple as possible. Brun ([2016](#)) describes Carnap's method of explication as "conceptual re-engineering." Carnap does not explain, however, how exactly these partially contradictory demands should be combined. In comparison, the proposed model of rational reconstruction contains clear normative and descriptive guidelines; and by distinguishing normal and revolutionary phases in the philosophy of science, it takes into account the possibility that incoherencies can arise between the two. There are also some similarities between rational reconstruction and the method of reflective equilibrium according to Goodman ([1955](#) , 85–89) and Rawls ([1971](#) , 20f, 48–52). However, while the method of reflective equilibrium works purely *coherently* , aiming at the reciprocal adaptation of methodological rules and intuition, rational reconstruction is *empirically grounded* by its descriptive corrective. Moreover, the normative corrective consists not of mere intuition, but of epistemological *hypotheses* , together with a supreme epistemological norm.

2.2 Common Epistemological Assumptions

The minimal epistemological model, common to more or less all scientific disciplines, contains five epistemological assumptions (E1–E5):

E1 – Minimal realism : According to this assumption, there is a reality, which exists independently of the (given) epistemological subject. It is not assumed that all properties of this reality are knowable. The possibility of fundamental limitations to scientific knowledge is left open and cannot be answered a priori, but only based on the factual epistemic success of science. Scientific disciplines aim to provide statements which are as true and as content-rich as possible about clearly defined areas of this reality. The term "truth" is used here in the sense of the *structural correspondence theory* , according to which the truth of a sentence consists in its structural agreement with the part of reality described by it. This notion of structural correspondence, which was first precisely defined by Alfred Tarski ([1936a](#)), does *not* imply any intrinsic relationship or mirroring between language and reality. It also allows that theories which are strictly speaking *false* may be true in an *approximative* sense, or *close* to the truth ([sec. 5.6.2](#) , 5.11.1).

E2 – Fallibilism and critical attitude : There is no infallible 'golden path' to the truth in the sense of the correspondence theory. According to the assumption of fallibilism, every scientific statement is more or less fallible; so, we can never be *absolutely* sure of their truth. Fallibility goes hand in hand with a critical attitude, according to which no statement may ever be once and for all exempted from criticism. However, we can regard the truth of a statement as more or less probable, and in the limiting case as "practically certain," in which case we consider its error probability negligible. For this reason, everything depends on using empirical tests to find out more about the *probability* of a scientific hypothesis, in the sense of assumption E4 below.

E3 – Objectivity and intersubjectivity : According to this assumption, a statement's truth must hold objectively, that is, independently of the beliefs and value-attitudes of the epistemological subject, because by assumption E1, reality has subject-independent existence, and truth is the correspondence between statement and reality. So, the assumption of objectivity already follows from assumption E1. However, the characterization of objectivity as subject-independence does not help us in the practice of science, as only subjects make statements and formulate hypotheses. We have no direct access to objective truth; we can only approach it indirectly via – not certain, but probable – *criteria* .

The second part of assumption E3 says that a central criterion for objectivity and also for truth lies in the *intersubjectivity* of scientific statements. Intersubjectivity, correctly understood, is a plausible consequence as well as a plausible indicator of objectivity and truth. It is a plausible consequence because if a statement's truth can be established at all convincingly, then – at least in principle – it must be possible to convince any person of the statement's truth, provided the person is sufficiently competent, that is, cognitively trained and informed. And it is a plausible indicator, because the agreement of many mutually *independent* and competent (i.e. fallible but truth-biased) researchers on the hypothesis provably increases its truth chances. This is an instance of the *Condorcet jury theorem* (cf. [Bovens/Hartmann 2003](#) , [sec. 5.2](#) ; [Dietrich/Spiekermann 2023](#)).

E4 – Minimal empiricism : With the exception of formal sciences such as mathematics, which we will discuss later, the objects examined by any science must in principle be accessible to experience and observation. After all, reliable information about reality can only be gained through perceptual observation; for perception is our only form of direct informational contact with reality. Empirical observations, then, are a decisive referee in the scientific search for truth: scientific law hypotheses and theories must be *tested* in terms of them. This is the central statement of assumption E4. Observations are not infallible (that would contradict assumption E2), but they offer the easiest and fastest route to intersubjectivity and practical certainty. Observations are “epistemically privileged,” as one also says. The empiricism contained in assumption E4 is *minimal*, because different from empirical reductionism it is not claimed that all scientific statements must be traceable back to observations via a chain of definitions, or must be provable by observations. Scientific theories can and should contain theoretical concepts, by which they can talk about things which lie beyond what is observable to human senses. The decisive requirement of E4 is merely that statements about the unobservable must have *observable consequences* by which they may be tested.

E5 – Logic in the wider sense : Using precise logical methods to introduce terms, to formulate sentences, and to construct correct arguments, is the most effective way to approach the goal of content-rich truth (according to assumptions E1–E4). For the meaning of a sentence cannot be established until the terms used in it have been clearly defined. And only sentences with a precisely formulated meaning may be examined for their logical consequences. Finally, only after the consequences of a hypothesis are known, may this hypothesis be precisely empirically tested in accordance with assumption E4. In conclusion, the procedure of empirical testing requires the use of exact logical methods at all times – in the wider sense of logic, not restricted to deductive logic.

Our minimal epistemological model makes as few assumptions as possible, but as many as necessary, in order to explain the possibility of objective science. Some of its assumptions are worthy of further explanation. We call the realism in assumption E1 “minimal” because it only asserts the objective *existence* of reality. Whether and to what extent reality is objectively discoverable cannot, according to assumption E2, be answered a priori, but only through the success of scientific discovery methods. This sets our minimal realism apart from metaphysical realism, in all its varieties. Moreover, we don’t claim that science is impossible without the assumption of a subject-independent reality, but only that this assumption is *normally* made in scientific disciplines (see [sec. 2.8.1](#)).

To avoid misunderstandings, we should add that the correspondence theory of truth, while well suited as a *definition* of truth, does not offer any *criteria* to evaluate the truth of a sentence (cf. [Rescher 1973](#) , ch. 1; [Kirkham 1992](#) , 195). The path to such criteria is illuminated by assumptions E5 and E4: investigate the empirical consequences of your hypothesis and test them using empirical observations. For example, to find out the truth of the sentence: “this flower is red,” the correspondence-theoretical insight that this sentence is true if the flower is red will not get me anywhere; rather, I must be in a position to apprehend the object denoted by the ostensive individual term “this flower” and the perceptual quality expressed by the predicate “red” visually, in order to test the truth of the sentence by a simple act of perception.

In assumptions E4 and E5, we claim the existence of only two minimal levels, in which discovery finds its origin and its touchstone: observation, which allows us fragmentary but reliable and intersubjective access to reality, and logic in its wider sense, which provides a universal cognitive tool, but in itself says nothing about reality.

In assumption E3, we limited our characterization of *intersubjectivity* to “sufficiently competent and informed” persons. This limitation is necessary, because persons who do not *understand* the content of a hypothesis, or who do not possess the relevant *evidence* , are unable to test a hypothesis for its truth. This difficulty shows us that intersubjectivity can on no account be used as a *definition* of objectivity or truth. The latter suggestion has been made in the constructivist school of Kamlah and Lorenzen ([1984](#) , 119), and one century earlier by the founder of the philosophical school of *pragmati(ci)sm* , Charles S. Peirce ([1878b](#)). But competence is an uncertain and gradual criterion, meaning that any number of persons judged to be competent may still err collectively. Peirce attempted to save his consensus definition by equating truth with the consensus that an *ideal* community of researchers with an infinite amount of time would approach in the limit. But even in this counterfactual version, an intersubjective consensus does not have to agree with the

truth of correspondence theory. There are countless questions on whose true answer even an ideal community of researchers can never achieve consensus, because there is no evidential access to them, although there is no reason to assume that these propositions have an indefinite truth value – for example, “How many microbes were located in Caesar’s nose as he crossed the Rubicon?” In conclusion, intersubjectivity can at most be an uncertain criterion for objectivity.

2.3 Common Methodological Features

From the supreme epistemic goal (G) and the five epistemological assumptions (E1–5), there follow very plausibly four methodological features (M1–4), which are common to all empirical sciences in the broad sense. First, a word on terminology. By an *observation sentence*, we understand a sentence that expresses an *observable* state of affairs. We distinguish between *actual* and *potential* observation sentences. We call an observation sentence actual if the state of affairs expressed by it has indeed *been observed*. By contrast, an observation sentence is merely potential, if the state of affairs expressed by it has not been observed, but could be observed, for example, at a later time or in a different place. Merely potential observation sentences express empirical *predictions*, for example, “there will be a solar eclipse tomorrow.” We can now formulate the four common methodological features as follows (similarly in [Weingartner 1978](#), [sec. 3.1](#) – [3.7](#)):

- M1* : Science searches for hypotheses which are as general and as content-rich as possible and recorded in a scientific language. In all disciplines, these hypotheses include *laws* and *theories*, and in some (e.g. the historical sciences), they also include *hypothetical singular sentences*.
- M2* : Science searches for *actual observation sentences*, as many as possible (and as relevant as possible), which reflect the results of observations, experiments, or measurements.
- M3* : Science attempts, with the help of general and hypothetical sentences, to *explain* the currently known actual observation sentences, and to *predict* new, and as yet unknown, *potential* observation sentences.
- M4* : Science attempts to *test empirically* its general and hypothetical sentences by comparing the predicted (potential) observation sentences with the currently known (actual) observation sentences. If the latter are in agreement with the former, the prediction was successful (it then becomes a successful explanation) and the hypothetical sentence (law or theory) is *confirmed* or *corroborated*. If there is a contradiction between the latter and the former, the prediction was without success and the law or theory is *falsified*, or in the case of a merely statistical prediction, *weakened*.

We only speak of an *explanation* if the state of affairs to be explained has already come to pass and is known, while in a *prediction*, the predicted state of affairs is new and unknown. Viewed logically, both explanations and predictions take the form of deductive or probabilistic arguments (see [sec. 6.1](#) – [6.4](#)).

These four methodological features or modules are closely connected with the epistemological assumptions of science. Feature M1 results from the epistemic goal G, because laws and theories are the typical examples of general, content-rich sentences. M2 results from assumption E4 (sharpened by assumptions E2 and E3): to discover more about the conjectured truth of its hypotheses, science must collect empirical data. In order for general hypotheses to be tested on the basis of these data, the logical or probabilistic consequences of the hypotheses, in the form of predictions or potential explanations, must be investigated, and these consequences compared with the data – in this way, the methodological building blocks M3 and M4 are motivated by the assumptions E4 and E5.

Successful predictions and explanations not only have a bearing on the ultimate scientific goal of finding true and content-rich theories, but they are also of crucial importance for human *planning behavior*, indeed, to our survival. Predictions allow people to plan their future by *adapting themselves* to the predicted sequence of events. Explanations are concerned with finding causes ([sec. 6.2](#)); they provide people with *knowledge of causes*. Causal knowledge gives people more power than merely predictive knowledge: it enables them to causally intervene in the course of events, and thereby *adapt* the world to suit *their own wishes*. The scientific subgoals of prediction and causal control have been pursued by humanity since the Stone Age. Anyone who can *predict* that nourishing plants are growing in a given area can decide to migrate there and thereby ensure a high reproduction rate for their own tribal group. But those who know the *causes* of the growth of nourishing plants, who know how to plant and cultivate them, need no longer be gatherers, but may become farmers, and practice agriculture instead.

In advanced science, the scientific procedures of explanation and prediction operate not on two but on three levels: actual observation sentences at the lowest level, empirical law hypotheses at the middle level, and scientific theories at the upper level. Between the first and the second, and between the second and the third level, the methodological procedures from M3 and M4 take place, as shown in [fig. 2.3-1](#). This threefold division of scientific work is advantageous for many reasons; in all advanced scientific disciplines, a division of labor has arisen between practical experimenters (levels 1 and 2) and theoreticians (levels 2 and 3). To give an example: In the case of the physical theory of the planetary orbits ([sec. 5.1](#)), level 1 corresponds to the actual astronomic observations (e.g. by Tycho Brahe), level 2 corresponds to the empirical laws of the (elliptical) planetary orbits around the Sun (inductively obtained by Johannes Kepler), and level 3 to Isaac Newton's theoretical explanations of these orbits by the Sun's gravitational attraction together with the planets' centrifugal force.

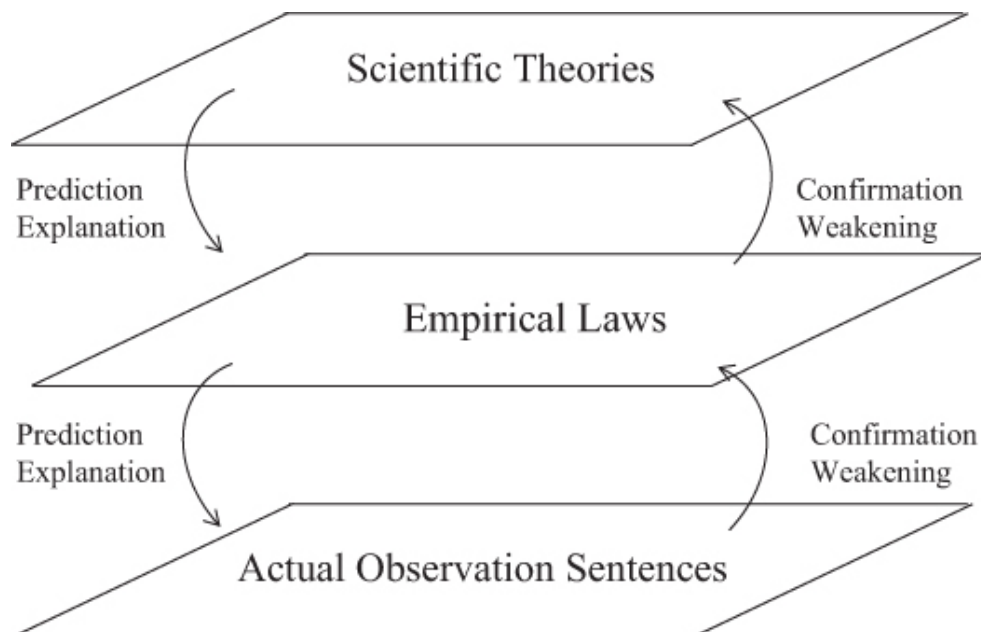


Figure 2.3-1: The three levels of the scientific method. [↗](#)

2.4 Scientific Disciplines and Their Classification

In this chapter, we will examine to what extent our epistemological assumptions E1–5 and the methodological features M1–4 match the existing scientific disciplines. To structure this investigation systematically, we introduce something that is rarely found in textbooks on the philosophy of science: a division of scientific disciplines according to their area of inquiry (see 2.4-1). Our division is certainly imperfect and we attribute to it not a final but merely a heuristic status. Some disciplines are mixed disciplines and are listed under “also;” further special disciplines can be added:

(2.4-1): *Scientific disciplines, classified according to area of inquiry :*

Sciences...

1. ... *of nature* : physics, chemistry, biology, geology, medicine (astronomy, cosmology, geography, paleontology, history of biological evolution)
2. ... *of technology* : mechanical and electrical engineering ..., also: computer science
3. ... *of human beings* : psychology (also: education, medicine, cognitive science)
4. ... *of (human) society* : sociology, economics, political science (also: anthropology, ethnology, geography)
5. ... *of (human) history* : history (also: anthropology, ethnology, philosophy as history of ideas)

6. ... *of human cultural (mental, social) artifacts* : legal sciences, linguistics, literary science, sciences of fine arts and music, art history, media studies (also: education, religious studies)
7. ... *of formal structures (formal sciences)* : mathematics (logic, statistics, theoretical computer science, system theory...), formal methodology and philosophy of science
8. ... *of the general foundations of human ideas* : philosophy (epistemology, philosophy of science and theoretical philosophy; ethics, esthetics, and practical philosophy)
9. ... *of God* : theology (also: religious studies)

Our classification has the advantage of not being tainted by partly value-based categories such as (natural) *science*, *humanities*, or *real science* (etc.). If we try to apply these catch-all categories to our area-of-inquiry-based classification, we recognize that there is no clear consensus about what goes where, and that different systems of classification often have more to do with the politics than the philosophy of science. Group 1, for example, covers the core *natural sciences*. But mathematics too sees itself as a natural science, although it belongs to the category of *formal sciences*; moreover, in the meantime, mathematics is not only applied in natural sciences but also in economics and social sciences. Of late, psychology also sees itself predominantly as a natural science, although it could equally be assigned to the humanities or social sciences.

Groups 5 and 6 comprise the core of the traditional *humanities*. Philosophy, though, is just as close to the natural sciences as it is to the humanities, although in contemporary universities, it is usually placed in the faculty of arts and humanities. In [sec. 1.3.8 – 1.3.9](#), we explained that since the 19th century, the traditional humanities were increasingly understood as a counterpart or even an opposition to the natural sciences. The term *humanities and social sciences* was coined in the 1960s and intended to serve as a modernized counterpart to the natural sciences. Whether this terminological group covers only groups 3 and 4, or also groups 5 and 6, was not terminologically fixed.

In the last two decades, the new terminological group of the *cultural sciences* was formed, in part as a reaction to contemporary tendencies in psychology and social sciences to increasingly use methods of natural sciences. The term “cultural sciences” is predominantly used in continental Europe and is understood in a broader sense than the Anglo-Saxon “cultural studies,” though its extension is not precisely defined. The frequently supposed opposition between natural and cultural sciences is an instance of what Snow ([1960](#)) has called the two ‘cultures’ in the current scientific age. In this book, we intend to refute the view that there exists a principal methodological *gap* between the sciences of nature and culture; this gap, where it exists, is more a matter of different value-attitudes and worldviews than a matter of unbridgeable differences in method.

In the term *factual sciences* (in German “Realwissenschaft”), all sciences are brought together that have a part of the *real* world as an area of inquiry; so, factual sciences include all groups apart from the formal sciences. Theology can only be seen as a factual science if we assume the reality of God, as this reality cannot be scientifically demonstrated, its inclusion here is questionable.

2.4.1 The Special Status of Mathematical Sciences

First, we have to explain the special status of group 7, mathematics, and formal sciences. In their totality, assumptions E1–5 and features M1–4 only make sense for the factual-empirical sciences. Mathematics and formal sciences do not have a domain of inquiry that consists of *concrete* and empirically accessible objects, but rather deal with formal and *abstract* structures. Mathematics, then, has no empirical basis, and therefore assumption E5 and methodological features M2 and M4, which require an empirical basis, do not apply. This does not mean that mathematics is a speculative discipline, which makes statements about reality with no empirical basis: pure mathematics makes no statements at all about an (empirically accessible) *reality*. All it does is develop formal structures and models, characterize them by using abstract axioms and definitions which are as general as possible, and derive logical consequences from them. It would be nonsense, for example, to say that the axiomatic system of real numbers or of Euclidean geometry was per se “true” or “false.” Such claims can only be made when mathematical terms are interpreted empirically, for example, when real locations or time points are assigned to real numbers, or real distances to mathematical vectors. But then geometry is no longer a mathematical theory, but becomes a physical one. Einstein once formulated it like this: “As far as the theorems of mathematics relate to reality, they are not certain; and as far as they are certain, they do not relate to reality” (cf. [Carnap 1967](#), 56). So, we can say: exactly *because* mathematics

ignores the empirical interpretation of its models, it is the strictest of all sciences. To be sure, mathematics concerns itself primarily with those structural models which are applicable to certain areas of reality. Seen in this way, mathematical disciplines can be more or less *fruitful* for the real sciences, but they cannot be true or false in the sense of the factual sciences.

Assumptions E2–4 and the methodological feature M1 can be applied to mathematics without restriction (the assumption of fallibility E2 insofar as even mathematical proofs are not guaranteed to be error-free). Assumption E1 of minimal realism applies only in an indirect way, insofar as mathematical structures have the purpose of being applied to the factual sciences. The methodological component M3 applies insofar as mathematicians are interested in the derivation of the various theorems, which express important properties of their models. Such derivations may be termed mathematical *explanations*, as long as it is clear that these are not causal explanations of real facts (cf. [Weingartner 1978](#), 102; [Pincock 2015](#); [Kuorikoski 2021](#)).

Let us not neglect the fact that the *philosophy of mathematics* includes opposing views on the nature of mathematical objects (cf. [Benacerraf/Putnam 1964](#); [Shapiro 2000](#); [Linnebo 2020](#)). According to the widespread view which is shared by us, mathematical objects are abstract *conceptual* objects, that is, conceptual constructions. Real objects, for example, five red roses, can *exemplify* mathematical objects, for example, the number five, but can never be identical with them. This view, called *conceptual structuralism*, was worked out in detail by the group of mathematicians around Bourbaki ([1961](#)) and philosophers of mathematics like Field ([1980](#)). The opposing standpoint in the philosophy of mathematics is *mathematical realism*, which assumes that mathematical objects exist *per se*. This position has been defended, for example, by Frege ([1918](#), 43) and Gödel ([1947](#), 271f). Also in this position, mathematical entities are seen as abstract and *platonic* objects, not located in the reality of space and time, but in a “platonic world of mathematical ideas” which is accessible not empirically, but only through mathematical “intuition.” An argument in favor of mathematical structuralism is Ockham’s razor (cf. 5.8.7): structuralism explains mathematical phenomena equally well as realism, but without the assumption of strange platonic entities. Another position is mathematical naturalism ([Maddy 1997](#)), according to which mathematical theories resemble scientific theories because they can fit reality more or less well in their applications. (The analogous view in the philosophy of logic has been called “anti-exceptionalism;” cf. [Hjortland 2017](#).) Against this view, it must be objected that in mathematical theories the reference to factual applications is fundamentally left open, whereas in empirical theories this reference is semantically fixed. As a result, empirical theories can be tested through experiments, which is undoubtedly *not* the case for mathematical theories.

In summary, mathematics and the other formal sciences are a *specialization* within the whole system of science, dealing with formal models, algorithms, and methods. This specialization has shown itself to be particularly *useful* in the history of science. Put in a nutshell, the information that the mathematician passes to the empirical scientist is this: “*if* your empirical domain satisfies the axioms of this-and-this mathematical structure, *then* it must also satisfy these-and-these theorems about this structure that we have proved; we grant you the satisfaction of the theorems, provided *you* grant us the satisfaction of the axioms.”

So, we characterize the formal sciences as *prerequisite and auxiliary sciences*, which furnish the factual sciences with formal foundations and structural models with known properties.

2.4.2 Testing the Common Foundations of Science

Having clarified the special status of mathematics, let us now ask: with the *exception* of the formal sciences, to what degree do the assumptions E1–5 and the methodological features M1–4 comply with the listed groups of disciplines? Obviously, E1–5 and M1–4 fit all disciplines in groups 1–3 (natural sciences, technical disciplines, psychology). There is also no reason not to apply them to the social sciences and historical disciplines in groups 4 and 5. Minimal realism, fallibilism, objectivity and intersubjectivity (E1–3), and logical clarity (E5) are accepted standards in all these disciplines, and the same is true for the search for general hypotheses, explanations, and predictions (M1, M3). In all these disciplines, empirical data are available, for example, in the form of the results of psychological experiments, interview data, economic and demographic data, and in the historical sciences in the form of historical texts and source materials, etc. By this, assumption E4 and building blocks M2 and M4 are also fulfilled. Of course, it is significantly more difficult to find general law hypotheses in social science and history, because of the complexity of their area of inquiry – but all that is required for the applicability of the methodological building blocks to a given

science is that the goals therein can reasonably be *sought after*, which does not necessarily mean that these goals are equally easy to attain in all different disciplines.

The same considerations apply to broad areas of the disciplines in group 6. Here as well, the epistemological assumptions E1–3 and (to a lesser extent) E5 are widely accepted standards. Moreover, these disciplines work with empirical data too; so, assumption E4 and the methodological building block M2 apply. For example, linguistic texts comprise the data for linguistics, and works of art and music from various cultures or stylistic periods are the data for literary criticism, art history, and music studies ([Kreuz/MacNealy 1996](#)). Members of the hermeneutic school raise the objection at this point, that “data” in the social sciences and humanities, for example, speech acts or texts, are always of an *interpretative* nature, so that understanding in the humanities is always subject to an unbreakable *hermeneutic circle* (cf. [Gadamer 1975](#) /2005, 235ff). But a number of philosophers of science have worked on the view that the testing of interpretative hypotheses concerning what an author intended to say proceeds in the same “holistic” manner as the testing of theories in the natural sciences (cf. Stegmüller [1979a](#) , 27–86; [Follesdal 1979](#) ; [Rosenberg 2008](#) , 43f). What hermeneutic philosophers rightly point out is that what one calls ‘data’ in the humanities are usually not raw data, but semantically interpreted data in the form of meaningful linguistic or nonlinguistic acts. However, also what physicists call data are usually theoretically interpreted measurement results, for example, elementary particles observed in the cloud chamber. This does *not* mean that raw or uninterpreted data or stimuli do not exist at all: in our physical example, they exist as condensation droplets in the cloud chamber; and in the interpretation of linguistic acts, they exist in the form of sequences of sounds or letters. Although some advocates of hermeneutics doubt the existence of raw linguistic data, their existence can easily be verified: you just have to hear someone speaking in a completely foreign language; what you will then hear are exactly these raw linguistic data.

The school of *methodological dualism* mentioned in [sec. 1.3.8](#) objects, once more, that there are no general law hypotheses in the humanities. If that were true, then the activities of these disciplines would have to be limited to methodological building block M2, that is to say, to the gathering of singular historical facts according to the “ideographic method.” What is true is that in the humanities and the social sciences, there are as good as no *strict* (or exceptionless) laws. But, as will be worked out later ([sec. 4.5](#) , [6.4.2](#)), there are *statistical* and in particular *normic* law hypotheses (“if A, then normally B”), which are indispensable for explanations and hypothesis testing in these disciplines. Therefore, the methodological components M1, M3, and M4 apply also to the historical sciences and humanities.

To avoid misunderstandings, when we say that the methodological components M1–M4 can be and have been successfully applied in the humanities, we don’t imply that *all* representatives of this group of disciplines make use of these methods. Within the humanities and social sciences, there is still the opposition between proponents and opponents of empirical-analytic methods – we mentioned Snow’s (1960) “two cultures” and Ross’s (1996) thesis of the “science wars” ([sec. 1.3.9](#)). Some scientists in the humanities and social sciences reject these methods simply *because* they are empirical and analytic. What this book nevertheless intends to show is that the methodological features M1–M4 are optimal means to pursue the supreme goal (G) of finding content-rich truths even in the human and social sciences.

There is, however, another phenomenon, found mainly in the humanities, that seems to stand in decisive opposition to the scientific method according to E1–5 and M1–4: namely, the phenomenon of *norms and value judgments* . We illustrate this problem at hand of a subdiscipline of legal sciences: jurisprudence, that is, the science of legal judgment. This discipline examines concrete actions and the facts of a case to determine how they fit into the given legal system, and specifically whether they fall under a certain category of offense or crime (going along with a certain punishment). The philosophy of *legal positivism* , which goes back to [Kelsen \(1960/2002\)](#), held the view that jurisprudence is a value-free systematic-deductive procedure, based only on *positive* law, that is, laws which have been passed by legislators. But Kelsen’s thesis is controversial in the philosophy of law. For example, was the removal of private property from a grocer’s shop *petty larceny* or *theft* ? (an example from [Savigny et al. 1976](#) , 101f). Or: does an underwear advertisement poster still conform to ‘public decency,’ or does it already violate ‘the dignity of women?’ (an example from [Hilgendorf 2000](#) , 15, 24). In these and other cases, it does not follow from the stipulations of the law and empirical knowledge of the facts, how the deed of the accused should be classified; and the judges and other interpreters of the law now have the task of *interpreting* the underdefined stipulations of that law. In [Savigny et al. \(1976\)](#) , 106, 144ff), it is shown that these interpretations are often based on *value assumptions* (or assumptions of justness), which cannot be found in legal texts, but rather come from an

‘intuitive legal consciousness.’ Now, a legal scientist can certainly show that *if* certain value assumptions are taken for granted, *then* a certain interpretation of the law is adequate. But how can she determine scientifically *which* value assumptions are justified? Can the legal scientist prove that a specified degree of misery of the accused person is sufficient for us to speak of ‘petty larceny’ instead of theft? Can she scientifically decide whether a billboard advertising underwear is still in accord with ‘common decency’ or is already an affront to the ‘dignity of women’? So the question which arises *not only* in jurisprudence, but also in other disciplines relating to the human or social sciences, is this: can fundamental value assumptions be scientifically justified?

Without further ado, the answer proposed by us is *no*. This does not mean that value- and norm-related issues do not occur in science. For one thing, the de facto values of persons can be *empirically* investigated. Moreover, specific value sentences can be logically deduced from *assumed* fundamental value sentences (and further premises). But a fundamental scientific justification of values is impossible. Our view here is concordant with the general consensus of (empirical) science and does not claim to be original; but our justification of this answer should contain some originality. For the value-neutrality requirement has been criticized so sharply by critics of empirical-analytical philosophy of science that it simply cannot nowadays be taken for granted, but needs to be worked through with great care. This task is addressed in [sec. 2.5](#).

2.4.3 Two Further Classifications of Sciences

The above subdivision of scientific disciplines was given according to area of inquiry. To end this section, we will explain two further possible systems of classification. One methodologically significant classification of the factual sciences follows the “invasiveness” of their empirical discovery methods. The classification is summarized in [fig. 2.4-1](#).

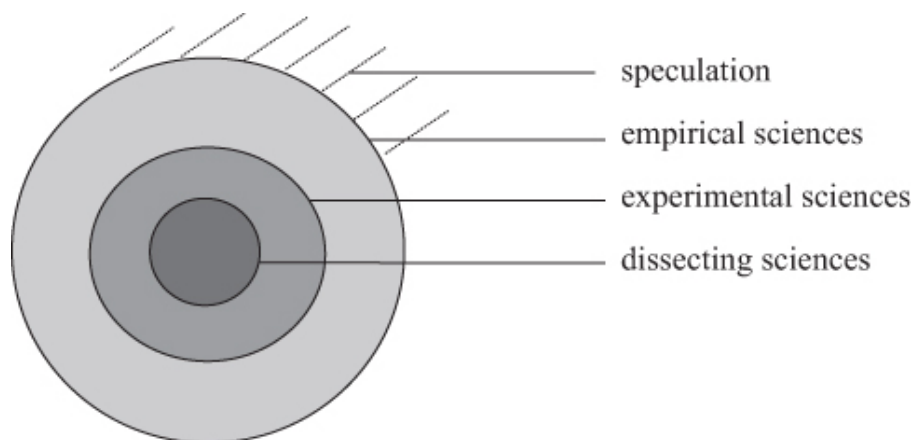


Figure 2.4-1 Classification of the factual sciences. [↗](#)

1. *Empirical sciences* (in the sense of minimal empiricism): By this, we understand all disciplines which refer in some way to empirical data, whether these data were obtained via free field observation, by studying historical records, or in laboratory experiments. To this most extensive group, we assign all disciplines described by assumptions E1–5 and methodological building blocks M1–4, including the humanities. Only those disciplines whose theories are based on pure *speculation* are not in this group – as for example mysticism, esotericism, speculative metaphysics, or the dogmatic and creed-based components of theology. Based on our epistemological assumptions, it is doubtful whether these are factual sciences at all.
2. *Experimental sciences*: This *subgroup* of empirical disciplines obtains its data from *controlled experiments*. The significance of controlled experiments as opposed to free field observation is that in an experiment, certain sets of conditions can be deliberately created, which give us information about the causal relevance or irrelevance of factors (see [sec. 4.2 – 4.4](#)), while field observers might have to wait ‘forever’ for a lucky coincidence to come up with the exact set of conditions required. For this reason, experimental sciences can explore their domain of inquiry much faster and more effectively

than sciences which rely on mere observation or passive reception of historical data. The experimental criterion separates the disciplines in groups 1–3 from the ones in groups 4–6, at least in tendency. For it is hardly possible, for both technical and ethical reasons, to perform experiments with large social groups, institutions, or legal systems. Of course, ethical concerns are already present in biological or psychological experiments, but here they are usually not strong enough to exclude them, only to restrict them.

3. *Dissecting sciences* : A significant subgroup of the experimental sciences are those which I will call – for want of a better word – the “dissecting” sciences. By this, I mean sciences, which not only expose their objects to various experimental conditions in order to analyze them intellectually, but which also *materially analyze* their objects, that is, take them apart or dissect them – perhaps in order to synthesize them afterward, that is, to put them back together again, provided this is still possible. I suggest that it is this dissecting character that constitutes the essential characteristic of *classical* natural sciences like physics, chemistry, biology, and medicine, in contrast say to experimental behavioral research or psychology.

Dissection of objects goes far beyond experimentation at the input-output level insofar as it makes the objects’ inner structure accessible to observation and experimentation. Chemistry and biochemistry have learned to break down items from our environment into their molecular and atomic constituents (analytical chemistry) and to put them back together again (synthetic chemistry). Nuclear physics has done the same at a *subatomic* level, here involving much greater amounts of energy. The analysis and synthesis of living creatures being worked on by science today – from genetic engineering and neuroengineering to artificial intelligence and robotics – could bring us to a point in the future where it is possible to genetically alter or even create a human or android being. Some people think these possibilities are terrifying; others that they are breathtaking. Here, as always, we should remember: scientific discoveries may be used for various purposes, both “good” and “evil.” But certainly, the possible risks and dangers going along with these technological possibilities are getting bigger and bigger. Therefore, the *ethics of science*, whose purpose is to set limits to the technological realization of the scientifically possible, is going to become more significant.

Independently of this, scientific disciplines may also be divided up according to the degree to which their concepts and *methodologies* are quantitative and mathematical, as follows:

Graduated division of scientific methodologies by increasing degree (←) of logical-mathematical and quantitative precision



Quantitative methods are on the left and *qualitative* methods on the right side of this spectrum. Note that on both sides we are looking at *scientific* methods – also qualitative hypotheses and methods can be pursued in a logically clean and empirically controlled way (cf. [sec. 3.1.4](#) ; [Denzin/Lincoln 2000](#)). The ideological polarization between quantitative and qualitative methods that is held by some qualitative researchers (e.g. [Lamnek 1988](#)) appears unnecessary and exaggerated. The strength of qualitative methods (e.g. case studies, narrative interviews) lies in advance of quantitative methods – in the exploration of relevant parameters and the generation of promising hypotheses. But a qualitative exploration has to be followed up by a quantitative-statistical analysis, as this is the only reliable way to test the generality of one’s hypothesis. That qualitative and quantitative methods are *complementary* in this sense is a widely held view among empirical researchers in the social sciences (cf. [Lazarsfeld/Barton 1951](#) ; [Giddens 1982](#) , 328); however, this view is not uncontroversial (cf. [Flyvbjerg 2006](#)).

2.5 The Question of Value Neutrality

2.5.1 Max Weber's Value-Freedom Postulate

The first detailed work on the value-neutrality requirement goes back to one of the main founders of sociology, namely, *Max Weber* (1864–1920). Weber was involved in a controversy with Gustav Schmoller, who propagated the view that social science should work out moral value judgments relating to a desired social order. Weber spoke out strongly against the pseudo-objectivity of such “podium values,” that is, value judgments presented with ‘scientific’ authority: Objectivity, said Weber, could only be reached in science if scientists limit themselves to factual, descriptive statements, and separate these clearly from their own values ([Weber 1922](#) /1949, ch. II and X).

The main reason for the self-imposed limitations of science to descriptive judgments, recognized by Weber, may be summarized thus: *Values are not properties inherent to objects themselves, but rather are based on subjective interpretations by us humans. Norms are also not objective facts, but man-made requirements, designed to aid the realization of certain values.* In the case of testing value and norm sentences, there is nothing which corresponds to an *empirical foundation* according to assumption E4 and methodological building blocks M2 and M4. At the same time, value and norm sentences are by no means purely formal claims, which might be justifiable by means of pure logic. There is therefore no objectivity in the sense of empirical science when it comes to norms and values. Ultimately, the *decision* for or against certain values or norms is a question of personal *freedom*.

This explanation of the value-freedom postulate is *prima facie* not only theoretically, but also practically plausible. Everyone knows from experience that finding a consensus on values between members of different cultures is much more difficult than finding a consensus on factual issues. Many strict Islamists would not take offense at Western science; the Western value system, however, seems twisted in their eyes. Nonetheless, the value-freedom requirement was massively criticized from the 1960s onward, especially by neo-Marxist movements such as the *critical theory*, and in the German-speaking area the Weber-Schmoller controversy flared up once more as the “positivism dispute” (cf. [Frisby 1972](#) ; [Adorno et al. 1976](#) ; [Dahms 1994](#)) and the “value judgment dispute” ([Albert/Topitsch 1971](#)).

Proponents of critical theory raised the objection that a purely descriptive or “positivist” method in social science had the effect of simply theoretically reproducing the status quo of oppressive social structures and thereby implicitly sanctioning them, while *critical science* must concern itself with *changing* oppressive social conditions ([Habermas 1966](#)).

This accusation is based on a fatal *misunderstanding* : descriptive research, of course, also includes research into natural laws, and thereby research into that which is naturally, technically, or practically *possible* and *impossible* . So, in fact, empirical social science is the very thing which can show us how social conditions may be *changed* , thereby contributing to the emancipation demanded by representatives of critical theory. In fact, almost all the “neopositivists” of the Vienna Circle had a *de facto* progressive political orientation and saw much of their work as being connected with social progress ([Dahms 1994](#) , 125; [Stadler 2001](#)). The only difference is that critical-empirical social science, as defended by Popper and Albert against Horkheimer and Adorno in the “positivism debate,” restricts itself to pointing out the *possibilities* for social change, and leaves normative decisions to the process of *democratic opinion-formation* , rather than wanting to justify these in a quasi-scientific way.

Even if the main accusation of critical theory was off the mark, the period that followed saw a number of legitimate objections being raised, which are in need of a more subtle discussion. Not only critical theorists (cf. [Dahms 1994](#) , 375), but also younger analytic philosophers warned that questions of values and norms should on no account be handed over to *irrationality* . Indeed, a rational treatment of value questions is both possible and necessary in society, and science has an important role to play in this regard. In this regard too, it is advisable to turn to Max Weber, who made subtle refinements to the value-neutrality requirement, in the face of just such objections:

First , according to Weber, it is absolutely possible for a social scientist to do research on the factual presence of value and norm systems, *without* passing judgment herself ([Weber 1922](#) /1949, 11f). To study polygamy in Islam, it is neither necessary to be an advocate nor to be an opponent of it. The statement that a

person or a society *believes* in certain values is not in itself a value statement, but a *descriptive* claim, whose confirmation requires only empirical data.

Second, also according to Weber ([1922](#) /1949, 21f), the scientist can identify *logical relationships* among value or norm sentences. Relevant to this task are the – meanwhile highly developed – disciplines of analytic ethics (e.g. [Frankena 1963](#)), deontic logic (e.g. [Aqvist 1984](#)), and decision theory (e.g. [Raiffa 1968](#)).

Third and most importantly, with the help of *means-end inferences*, the scientist can infer from *given* norms and descriptive knowledge a series of derived norms about optimal means of reaching accepted ends and pass these on as recommendations to society and politics. Because this represents the most important function of science for society, the next section is devoted to it.

2.5.2 Value Neutrality and Means-End Inferences

Finding suitable means for given ends is an important task of descriptive science. It obeys the following logical form:

(Def. 2.5-1) *The general scheme of means-end inferences* :

Descriptive means-end premise : M is in the given circumstances C a necessary – or alternatively an optimal* – means for the realization of end E.

Thus – hypothetical value- resp. norm conclusion : If end E is accepted as a fundamental value (or norm), then the realization of means M is a derived value (or norm).

Means-end inferences are accepted by most ethical theories as analytically *valid* (the asterisk* is explained below). The conclusion can be formulated for value statements (“P is or is not valuable”) and normative statements (“P ought or ought not to be done”). Both are analytically related, which can roughly be expressed by the formula: “The good ought to be done and the bad ought to be avoided” (for details, see [Schurz 1997](#) sec. 11.2–11.3). Value judgments cannot be analytically reduced to factual statements about the value sentiments of some or many people (as [Anderson 2004](#) proposed), but rather possess *normative force*.

The recommendation of means in the conclusion is not understood in a categorical, but in a hypothetical or conditional sense. That is, it is relativized to the assumed ‘fundamental’ or initial values (ends) E: *If* such-and-such fundamental values are assumed, *then* certain action recommendations follow as derived instrumental values. For conceptual clarification: A *categorical* value statement has the form of a basic value sentence “P is (or is not) valuable;” a *hypothetical* or conditional value statement takes the form of an if-then sentence “If P is (not) valuable, then Q is (not) valuable” (this distinction goes back to [Kant 1785](#)). A categorical value statement is *fundamental* if it cannot be derived from other categorical value statements using means-end inferences; otherwise, it is *derived*.

It is crucial for the value neutrality of science that recommendations of means are formulated hypothetically rather than categorically. Only in this way can a recommendation of means follow from factual knowledge and thus remain value-neutral itself, and only in this way can objectively value-neutral science be reconciled with societal value engagement. Value judgments derived hypothetically by science are also referred to as *instrumental* rationality; however, we avoid this term, as it implies a certain devaluation (“merely instrumental”), whereas what we mean is a genuine value engagement – only formulated hypothetically in order to clearly *separate* the factual component from the value component of the statement.

Another advantage is that, through the use of rich scientific knowledge, a large number of derived norms can be obtained from just a few fundamental norms. For example, the EU climate goal of becoming carbon-neutral by 2050 entails a multitude of recommendations of means. In this way, the burden of justification can be shifted – if not entirely, then to a large extent – from the normative to the descriptive.

In a means-end inference, it is assumed that M is either a necessary means – that is, every realization of the goal requires the realization of M – or an optimal means, meaning that among all practically possible alternatives, M maximizes total utility with regard to goal realization and side effects. Strictly speaking, the definition of an optimal means requires a listing and cost-benefit evaluation of all alternative means. In order for the hypothetical recommendation of optimal means to remain value-neutral, this cost-benefit evaluation must be included as a fundamental value assumption in the antecedent (the “if”-part) of the recommendation

– which is indicated by the asterisk. The precise reconstruction is therefore more complex than presented in [def. 2.5-1](#) (see [Schurz 2024b](#), ch. 3).

If M, on the other hand, is a merely *sufficient* means for an end E, then the means-end inference is not valid, because many sufficient means cause harm that outweighs the benefit of achieving the end. For example, to achieve the end of letting fresh air into a room, smashing a hole through the wall is a sufficient but not recommendable means; by contrast, opening a window is an optimal means, and a necessary means is some sort of opening to the outdoors.

Neither the inference to necessary nor to optimal means implies the mistaken thesis that “the end justifies the means:” if the costs of every possible realization of the end E exceed its benefits, then E should not be realized and can no longer be considered an acceptable fundamental value (given a consequentialist understanding of value).

The scientists or experts should take the fundamental norms from the knowledge *users* – that is, the public or their political representatives. Based on these norms, the scientists can then derive recommendations of means from scientific knowledge and pass them back to the knowledge users. Let us illustrate this by means of educational *teaching theories* (cf. [Flanders 1970](#) ; [Patry 1991](#)). Let us take it that *cognitive success* (i.e. mastery of material to be learned) is the fundamental goal of teaching. Based on empirical research, an education scientist may reach the descriptive hypothesis that cognitive success may be maximized if between 70% and 80% of teaching time is teacher-guided. The education scientist passes on this conclusion to teachers as a *derived* norm. This procedure is, however, only in accordance with the value-neutrality requirement if the means recommendation is formulated *hypothetically* : if your goal is to maximize cognitive success, then 70%–80% of class time should be teacher- guided.

If the explicit reference to assumed values is neglected, ideologically worrying consequences can result. Assume that in our example, the teachers have different fundamental goals than the scientists; they see *social learning* as an equally important goal to cognitive success. Let it be empirically shown, perhaps even by the same scientist, that the optimal means to reach this purpose is a maximum of just 60% of teacher-guided class time, alongside 40% of group activity. What if the scientist in our example, however, passes on her recommendation categorically rather than hypothetically, for example, she asserts *80% of class time should be teacher-guided* , and teachers believe him? Now the scientist is *manipulating* the teachers, because she is concealing the fact that her means recommendation is valid only for *her own* value assumptions. The teachers, who *do not* share these values, are being badly advised on how to reach their own chosen ends.

The same is true of many other situations, from medicine (a doctor advises a patient) to law (a lawyer advises a married couple about to get divorced). It is highly important in all such cases that the scientific “expert” makes the value assumptions *explicit* that she assumes in her means recommendations, as this is the only way that the “customer” being advised can find out whether these correspond to his own values. With these examples, we *emphatically* point out that the value-neutrality requirement does not have a dogmatic, but rather a *democratic* and *emancipatory* function. In [sec. 2.10](#) , we will defend this thesis in the context of contemporary debates on values and their role in science.

In summary, the central conclusion of this section is that value neutrality and value engagement in science are indeed compatible. The oversimplification of the value-neutrality requirement into the “value-free doctrine,” according to which values have “no place” in the empirical sciences, is therefore a serious misrepresentation that often serves as an easily criticized straw man for opponents of value neutrality. Sciences can and should use means-end reasoning to derive values from externally given fundamental values and pass them on to knowledge users in the form of hypothetical recommendations of means. Only the *justification of fundamental values* lies outside the scope of objective science.

2.5.3 Explication of the Value-Neutrality Requirement

To reach the final formulation of our value-neutrality requirement, we need to take two further steps of clarification.

1. Many authors have argued that science cannot be value free in an absolute sense, because the whole enterprise of “science” itself rests on certain goals and values: the so-called *science-internal* or epistemic values (e.g. [Schmidt 1959](#) , 648ff; [Doppelt 2007](#) , 190ff; [Sober 2007](#) , 110). In [sec. 2.1](#) , we identified the search for general and content-rich truths as the supreme epistemic goal of all sciences.

This fundamental internal value of science is accepted by all scientific disciplines, from which a range of derived epistemic values follow as means to achieve this supreme epistemic value (see [Doppelt 2007](#)) – in particular, the four methodological characteristics mentioned in [sec. 2.3](#). The value-neutrality thesis must therefore be restricted to fundamental *science-external* or non-epistemic value assumptions.

2. Is the restriction to fundamental science-external values and norms enough to give us a sustainable version of the value-neutrality requirement? *No* – insofar as one views science not merely as a process of justification, but as a *social process*. The process of scientific discovery is usually divided into three phases: the *context of origination* (CO) (also called “context of discovery”), in which the scientifically relevant problems are first identified and defined, the *context of justification* (CJ), in which data are gathered and systems of hypotheses generated and tested, and the *context of application* (CA), in which well-tested discoveries are used for various different social purposes. In both the CO and the CA, non-epistemic values inevitably play a role. Decisions about which research questions are considered important enough to pursue are also influenced by non-epistemic interests – such as those of funding bodies or the research team. Even more strongly, non-epistemic interests determine the purposes for which scientific knowledge is used in the CA. What the value-neutrality thesis demands relates exclusively to the CJ: in CJ, science-external values must not play any role. In conclusion, we can formulate and schematically represent the value-neutrality requirement as follows (see [fig. 2.5-1](#)):

(Prop. 2.5-1) *Value-neutrality requirement (V)* : A specific realm of scientific activity, namely its context of justification, should be free from fundamental science-external value assumptions.

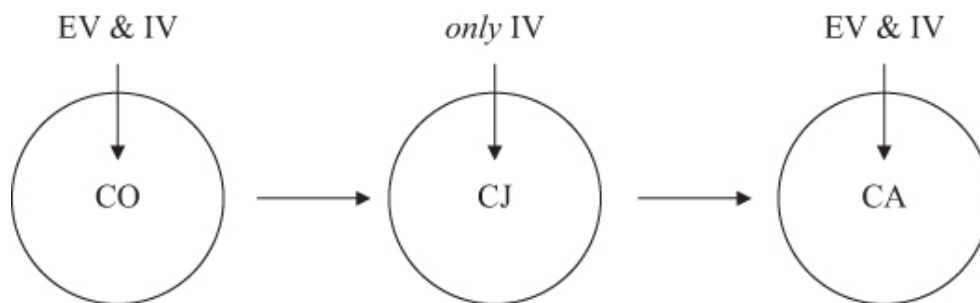


Figure 2.5-1 Schematic representation of the value-neutrality requirement. [↗](#)

Legend: CO – context of origination, CJ – context of justification, CA – context of application; IV – science-internal values, EV – science-external values.

The value-neutrality requirement will hereafter be abbreviated as (V). Note that (V) is a normative requirement, not a factual claim about the actual judgment practices of all or most scientists. There is broad acceptance of the value-neutrality requirement in the sciences, but as noted in [sec. 1.2.9](#) and [2.4.2](#), the requirement is rejected by a significant number of scholars in the humanities and cultural studies (keyword “science wars”). The content of the value-neutrality requirement can also be expressed as a thesis without an explicit normative formulation as follows: the *justificatory status* of scientific claims is *independent* of non-epistemic values. However, even in this formulation, it remains a statement about the epistemic-normative justification status of scientific judgments – and not about the actual judgment practices of scientists.

In our modern, value-pluralistic information societies, politicians and the media rely on scientific legitimations. As a result, scientists face a constant pressure, or at least the temptation, to present recommendations of means in the form of categorical value judgments. Scientists should resist this temptation, as doing so ultimately undermines public trust in science, as recent empirical studies have shown (see [Angeli et al. 2021](#); [Intemann 2024](#), 10). The distinction between factual and value judgments is an essential prerequisite for any credible differentiation between true facts and fake facts, and thus for countering science denial and conspiracy thinking in our time. Only objectivity and value neutrality can enable stable scientific progress through the politically and culturally shifting phases of history.

2.6 The Demarcation Problem and the Methodological Unity of Sciences

2.6.1 Demarcation

If we add the value-neutrality requirement (V) to the epistemic goal (G), the epistemological assumptions (E1–5) and methodological features (M1–4), we can thereby characterize the boundaries of the scientific. As long as legal science, for example, analyzes empirically occurring legal systems, or establishes logical and means-end relationships between norms, it remains a descriptive science. The logical analysis of ethical statements carried out by analytic ethics and meta-ethics also fits easily into the logical-descriptive scientific program. Literary and art theory, which deal with the aesthetic effects of various literary and artistic genres, fit into the framework of empirical-descriptive science, as long as aesthetic effect refers to a *factual* effect on people. Only absolute statements about aesthetic values, which are not relativized to a factually given artistic style, go beyond the scope of descriptive science. This criticism also applies to the tendencies within humanistic psychology and pedagogy which place values like personal development and dignity above objective analysis ([Bühler/Allen 1973](#)). Religious studies, as long as they reconstruct factually occurring religious systems, belong to the descriptive scientific program, while theologies that dogmatically assume the truth of a certain religious, “creed” are departing doubly from the realm of science: first, by assuming the truth of speculations about the nature and existence of God and second by making fundamental value assumptions.

In summary, we can demarcate the boundaries of science as follows: *Sciences in the sense used here are all empirical disciplines in the broad sense (according to G, E1–5, M1–4 and V), together with their formal auxiliary sciences and their means-end-based applied disciplines.*

This characterization of “science” gives us a demarcation between scientific and nonscientific disciplines and fields of mental activity. So, it serves as a suggested solution of the *demarcation problem* (cf. [sec. 1.1](#)). We do not understand this demarcation problem, though, in the same way as logical positivism or critical rationalism did. *First* , we do not identify the nonscientific with the cognitively nonsensical (as e.g. [Carnap 1928b](#)). This identification is unacceptable, because empirically vacuous sentences like “God exists” or fundamental norm sentences like “Help your neighbor” may of course be *understood* , and so have a communicable *meaning* . *Second* , we do not devalue such sentences; we classify them as nonscientific, simply because there is no way to test them empirically. *Third* , we see the significance of the demarcation problem differently from Carnap. Our proposed solution of the demarcation problem does not pursue the inner-philosophical concern of eliminating a part of philosophy called “metaphysics.” Rather, it serves (A) a societal demarcation aim and (B) an evolutionary explanatory concern:

Point A concerns the public debate on the question which kinds of information are sufficiently objective and confirmed to be taught in public schools, and which are not. This requires demarcation criteria between science and religious ideology. As an example, we mentioned in [sec. 1.1](#) two famous cases in which judges based their decisions against the teaching of creationism in public schools on demarcation criteria developed by philosophers of science. However, there is a controversy among philosophers of science themselves as to what the right demarcation criteria should be. Ladyman and Ross ([2007](#) , 30) prefer sociological over logical demarcation criteria. But sociological demarcation criteria are insufficient. For example, Pennock ([2011](#) , 193ff) reports that Judge Jones suggested the application of peer-review methods as a social demarcation criterion. The reaction of Behe, a prominent intelligent design defender, was to promote peer reviews within the creationist camp. Behe pointed to a couple of creationist writers that indeed had been peer-reviewed – not by naturalistic scientists, but by other intellectual creationists. But the consent which a creationist writing earns from other creationists is clearly not an indication of its scientific character.

Point B concerns the *explanation* of the *evolutionary impact* of the sciences. Looking back to the last one or two millennia of human development, science has made extremely rapid progress, compared to ethics, religion, and arts which have developed slowly, at least in those domains where they have not been driven by external technological change. The factual differences in the level of evolutionary impact of the sciences in comparison to ethics, religion or art, and the degree in which they have changed our life, are *so* marked, that they positively beg to be explained. Anyone who wants to understand the powerful independent dynamics of

science and technology has to find out, *what it is* that separates science so strongly from nonscientific systems of ideas. In conclusion: Any philosophy of science which wants to be of practical significance must face up to the demarcation problem.

2.6.2 Methodological Unity instead of Physical Reductionism

Our demarcation account also contributes to the question of the *unity of science*. However, our understanding of the unity of sciences is not ontological, but methodological in nature, and therefore different from traditional accounts. These accounts were predominantly based on *reduction* attempts, that is, on attempts to provide an explanatory or ontological reduction from 'higher-level' or 'special' sciences to *physics*. There are, however, weaker and stronger sorts of reduction accounts. The career of 20th-century reduction accounts started with the program of *physicalism* of Otto Neurath, a prominent logical empiricist, who announced this program as a means to establish the unity of sciences. Neurath's physicalism was weaker than it sounds, insofar, it consisted only in the requirement that all scientific sentences should be translatable into statements about concrete spatiotemporal things and properties ([Neurath 1934](#) ; [Dahms 1994](#) , 170).

A much stronger form of reductionism had been developed in the 1950s and 1960s: the program of *physical micro-reduction*, whose goal was an ontological or explanatory reduction of the laws of special sciences to laws about underlying physical microstructure. ¹

There exist different concepts of reduction. In the standard account of Hempel ([1969](#)) and Nagel ([1970](#) , 125), a physical reduction of a law describing the behavior of a complex system consists in the derivation of this law, or an approximative version of it, from fundamental laws of physics together with contingent boundary conditions, and suitable correspondence principles that translate higher-level concepts into physical concepts. More recent notions of reduction are discussed in van Riel and van Gulick (2025); as these discussions suggest, the core ideas of the Hempel-Nagel concept of reduction can still be considered well-established.

In spite of some positive examples (e.g. the micro-reduction of temperature to mean kinetic energy), this reduction program was not especially successful. It does not describe the way in which special sciences typically proceed. The objects of the special sciences are more-or-less complex systems. At every level of (at least moderately high) complexity, special sciences have discovered their own laws, which are called *system laws* in [sec. 6.5.1](#). Examples are the chemical laws describing the behavior of molecules, biological laws about the functioning of organisms, psychological laws describing human behavior, etc. There exist three arguments, which show that the laws of special sciences are typically *not* reducible to physics, at least not fully (cf. [Cartwright 1983](#) , 104f, 113f):

1. The argument from *hypercomplexity* : The fundamental physical law governing a specific physical system is expressed by its Schroedinger equation. However, already for moderately complex microsystems such as a medium-sized molecule (as opposed to a simple atom), their Schroedinger equation is far too complex to be solvable without radical simplification assumptions. These simplification assumptions do *not* come from fundamental physics but from empirical chemical knowledge. For these reasons, most chemists are convinced that chemistry will never be *fully* reducible to physics (cf. [Hendry 2008](#) , 525ff.; [Mortimer 1986](#)). Much more is this true for the laws that describe highly complex systems such as biological systems: here already the formulation of an approximately complete Schroedinger equation is impossible and attempts to derive biological normal-case laws ([sec. 4.5.2](#)) from fundamental physics and factual descriptions of the physical microstructure of an organism are hopeless. Rather, biological laws are typically discovered by empirical-inductive means. This does not exclude the existence of *partial* reduction relations, for example, between biochemistry and biology or neuroscience.
2. The argument from *multiple realizations* : Empirically, identical macro- properties of a complex system, such as elasticity or intelligence, can be realized by very different physical microsystems ([Fodor 1974](#)). It is hard to find general principles that translate a given macro-property into an exhaustive disjunction of physical micro-properties.
3. The argument from the *systemic properties* : The parts of a complex system are systemic, which means that the properties of these parts in isolation or in a simpler system are different from the properties of the parts within the complex system. An example are the properties of the hydrogen atom H in water

molecules H_2O (O for oxygen), which are different from the properties of the H-atom in isolation, or in the H_2 -molecule. ²

The argument from systemic properties only shows that a so-called horizontal reduction is usually impossible, that is, the explanation of a complex system by the properties of its parts in isolation or in simpler systems. By contrast, “vertical reduction” means that the higher-level properties of a system are reducible to the lower-level properties of its parts *within* the system. Boogerd et al. ([2005](#)) argue in application to biochemical networks that in spite of horizontal irreducibility, vertical reducibility is often possible, and is sufficient for explanations based on biochemical mechanisms. However, since the systemic properties of the parts are determined by the entire structure of the complex system, a micro-reduction of the systemic properties of the parts would lead into an explanatory circle and is therefore not feasible.

These arguments show that the hopes for a physical micro-reduction of the laws of complex systems are dim. For this reason, many philosophers of science have replaced the idea of explanatory reduction by the weaker requirement of *supervenience* ([Hüttemann 2004](#)). Supervenience requires not an explanatory reduction, but merely a “vertical” ontological determination: the physical microstructure of a system *determines* its higher-level properties (at the same given time point). As this determination relation may be infinitely complex, it does not imply the possibility of a reduction by human cognitive resources.

The opposite of supervenience is ontological *dualism*, according to which two physically identical systems may nevertheless differ in some higher-level non-physical properties, for example, “having a mind” according to mind-body dualism. In fact, the supervenience position has been developed by Davidson ([1970](#)) as a naturalistic counter-position to Cartesian mind-body dualism. The supervenience position is fully compatible with the possibility, for which we have argued in this section, that an explanatory reduction of the sciences of complex systems to physics will not be possible in the foreseeable future.

In conclusion, the prospects to establish the unity of sciences by methods of a reduction to fundamental physics are not promising. What one finds along this route is not unity, but *ontological variety*. We suggest that the unity of the sciences should rather be sought by highlighting their epistemic-methodological common ground. The unity of the sciences, if it exists at all, does not exist at the ontological level, but in the form of their *methodological unity*. Of course, there are also important methodological differences between certain disciplines, which will be worked out in several sections of this book ([sec. 1.3](#), 2.4, 3.1, 4.3, 4.5, 5.1, 5.3, 5.5, 6.4, 6.6). But our considerations so far have shown that there exist essential methodological features which are common to all empirical disciplines, despite their differences, and which delineate scientific from nonscientific intellectual activities: the ultimate epistemic goal G, the epistemological assumptions E1–5, the methodological building blocks M1–4, and the value-neutrality requirement V. In this methodological version, the idea of the unity of science can live on in a humbler, but more sustainable version.

2.7 Scientific Inference and Argumentation

In the methodological building blocks M3 and M4, scientific processes of argumentation and inference take place. Thereby, we are dealing with very different processes of inference, which are explained in this section.

2.7.1 Deduction and Induction

When, in accordance with M3, we use a strict law hypothesis together with observation sentences to infer further observation sentences, then we are dealing with a logical-deductive inference. But when, in accordance with M4, we infer from the fact that a large number of actual observation sentences are instances of a law hypothesis, that the law hypothesis is likely to be true, then we have made an inductive inference, or more accurately an inductive generalization. Here are two examples:

(2.7-1)

Deductive inference (logic i.n.s.):

All fish breathe using gills.

Inductive generalization inference:

All fish observed so far (no. 1, 2, ... n)

This animal is a fish.	have breathed using gills.
Therefore, this animal breathes using gills.	Therefore (probably), all fish breathe using gills.
<i>Certain</i> : truth transfer in all possible worlds.	<i>Uncertain</i> : truth transfer only in sufficiently “uniform” possible worlds.

The decisive difference lies in the fact that only deductive inferences transfer the truth from the premises to the conclusion with *certainty*, that is, in all possible “worlds:” *if* the premises (the sentences above the dividing line) are true, then the conclusion (the sentence beneath the dividing line) is also certainly true. Deductive inferences constitute the object of *logic in the narrow sense*. (Some authors place an inductive logic in opposition to deductive logic, for example, [Carnap 1950b](#), sec.43ff; this terminology is disputed, however, see [sec. 5.9.5](#).)

By contrast, inductive generalizations are fundamentally uncertain – the double line in ([2.7-1](#)) indicates this uncertainty. This is because the premises of an inductive generalization tell us only about the cases observed *so far*, while the conclusion generalizes to *all* and in particular to all *future* cases. For this reason, it is also said that inductive inferences are *content expanding*. Only in circumstances or worlds which are sufficiently *uniform*, whose future and past are sufficiently similar, can we reliably infer the truth of an inductive conclusion from the truth of the premises. Nothing can logically guarantee that the future course of events will be sufficiently similar to the course of events observed so far. An extinct fish without gills could, for example, be discovered tomorrow, or a new species of fish without gills could start to evolve through mutation.

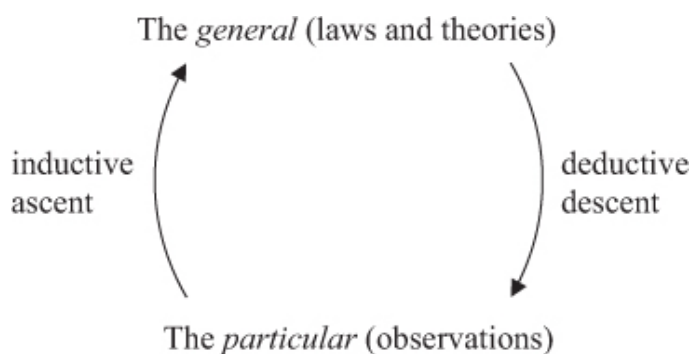


Figure 2.7-1 The inductive-deductive scheme. ↵

This was (very roughly) the main argument of *David Hume*, who was the first philosopher to show that a justification of inductive inference is impossible by means of deductive logic. Hume went on, however, to cast doubt on the rational justifiability of induction as a whole. Hume’s “problem of induction” belongs to the most difficult philosophical problems; a closer look at its subtleties will be given in [sec. 2.11](#).

We have so far explained induction and deduction in the sense of the so-called *inductive-deductive scheme*. According to this scheme, which was already outlined by Aristotle, induction is a method of reasoning (or ascending) from the *particular* to the *general*, and deduction is a method of inferring (or descending) from the general back to the particular (see [fig. 2.7-1](#)).

Although this scheme includes some important methodological features of the sciences, it is *oversimplified* or even incorrect in several respects. For one thing, the binary division of types of inferences on which it rests is *too narrow* – there are far more types of inference, which we will briefly list here. Where the general is concerned, we must distinguish between *strict* (or deterministic) and *statistical* generalizations.

1. Types of deductive inferences :

1.1 From the (strictly) general to the particular – for example, All Fs are Gs, this is an F, therefore: this is a G.

1.2 From the general to the general – for example, All Fs are Gs, All Gs are Hs, therefore: all Fs are Hs.

1.3 From the particular to the particular – for example, This is an F or a G, this is not an F, therefore: this is a G.

By contrast, there are no deductive inferences from the particular to the (essentially) general.

2. *Types of inductive inferences* (see also [Carnap 1950b](#), 207f):

2.1 *From the particular to the general* :

2.1.1 *The strict inductive generalization* : All Fs observed so far were G, therefore (probably), all Fs are Gs.

2.1.2 *The inductive-statistical generalization* : r % of all Fs observed so far were G, therefore (probably), approximately r % of all Fs are Gs.

2.2 *From the particular to the particular – the inductive prediction* : All Fs observed so far were G, therefore, the *next* F will probably be a G, too.

2.3 *From the general to the particular – the inductive-statistical specialization* : r % of all Fs are Gs, this individual is an F, therefore, we can be r % sure that this individual is a G.

Inductive specializations are also called “direct inference” ([Levi 1977](#)) or “statistical principal principle” ([sec. 3.13.4](#)). Two facts about them are remarkable: First, the strict variant of inductive specializations is deductively valid. Second, like inductive generalizations, inductive-statistical specializations assume uniformity: they are only reliable if our actual observation sample is representative for the total domain.

2.7.2 Popper and the Significance of Inductive Inference in the Sciences

The founder of critical rationalism, Karl Popper, advocated the thesis that the induction problem is, on the one hand, unsolvable, but, on the other hand, that empirical science can proceed without inductive inferences altogether (1935/2002, sec. I). In the following, we want to demonstrate, based on a differentiation between *three kinds of induction*, that Popper’s arguments bring to light important insights, but that his radical claim that empirical science could go without induction is not tenable (cf. [Schurz 2002a](#)).

1 *Methodological induction*. This view of induction connects directly with the inductive-deductive scheme explained in [sec. 2.7.1](#) . According to this view, induction is a method of inferring general laws and theories by inductive generalization from particular observations. Popper criticizes the view contained in methodological induction that induction is the central method of discovering or generating law hypotheses and theories. The belief that such a discovery method is at all necessary in science rests, says Popper, on a confusion of the *context of discovery* and the *context of justification* (1935/2002, sec. I.1–2). How scientific hypotheses are generated, be it through induction, intuition, or trial and error, is completely *irrelevant* to the context of justification. Seen like this, methodological induction is *dispensable* in the context of the justification of science.

By contrast, call to mind [sec. 2.1](#) : not all but *some* methods of generating hypotheses may give us a certain measure of justification as well. This is the case, for example, for the inductive generalization inferences explained in [sec. 2.7.1](#) : these inductive inferences are methods of discovering empirical law hypotheses, which at the same time inductively justify these law hypotheses, that is, preliminarily confirm them. More refined inductive discovery algorithms can be found in Langley et al. ([1987](#)) (see also [sec. 4.2.2](#) , 4.3.2). Hypothesis generating procedures which at the same time give us a justification are of course cognitively much more *efficient* than blind trial and error procedures.

In two important points, however, Popper’s criticism of methodological induction is right. (i) Hypotheses do not *have* to be produced by methodological induction – all methods of producing hypotheses are allowed and deserve critical assessment. (ii) Inductive generalization methods work for empirical law hypotheses, but they are no longer sufficient for the construction of scientific theories. Even very simple theories – like, for example, the Galilean law of free fall, which posits the idealization of a friction-free fall – cannot be created by inductive generalization procedures alone, since they contain concepts that transcend the observable – so-called *theoretical concepts*, such as “friction-free movement,” “mass,” or “force.” So Popper’s criticism of

the possibility to create scientific theories by methodological induction (see 1983, 118) is to a large extent legitimate.

2 Logical induction . This view of induction goes back to Carnap ([1950b](#)) and understands induction not as a method of discovery, but as one of justification of scientific hypotheses, that is, as a method of determining the *degree of confirmation* conferred on theories by given observational data. This should happen by the establishment of a system of logical probability, in which the inductive probability of a theory given a body of observational evidence may be determined in a rational way (see [sec. 4.7](#)). Instead of the inductive degree of confirmation, one can alternatively regard the degree of (epistemic) *truthlikeness* of a theory relative to a body of evidence as the central measure of evaluating theories (see [sec. 5.11.1](#)).

Popper's criticism of the view of logical induction now says the following (1983, 19f, 23): while we are able to determine that relative to given observations, a theory T_1 is *better* confirmed or corroborated than another theory T_2 , or that T_1 is closer to the truth than T_2 , we still cannot give an *absolute* value for the probability or truthlikeness of T_1 or T_2 , as tomorrow a new superior theory T_3 may be developed, which completely overshadows both T_1 and T_2 . Popper's argument represents a substantial insight into the *limitations* of logical induction. In order to assign a nonarbitrary *numeric value* to the inductive probability or truthlikeness of a theory T given some evidence, one must have at one's disposal, among other things, a comprehensive list of all relevant *alternative theories* $\{T_1, T_2, \dots\}$. However, to establish such a list in a nonarbitrary manner is, at least in the case of theories, impossible. For any given theory T , there are always indefinitely many alternative theories, as these may contain any number of new theoretical concepts. A measurable linguistic space containing all possible alternative theories simply does not exist (see [Howson 2000](#) , 46). Many theories that were later accepted were still unimaginable a century earlier in the history of science (see [Stanford 2006](#)).

Again, in the case of *empirical* hypotheses, the situation is different. Here, it is possible to establish a finite list of all possible alternative hypotheses in a *limited* observational vocabulary. Let us consider, for example, the set of all interval hypotheses of the form $H_n =$ "between n and $n+1\%$ of all people are brown-eyed," with $0 \leq n \leq 99$, and let us assume a uniform distribution of the initial probabilities of these hypotheses. Then, the inductive probability of H_n , given a sample of say m people, including $k \leq m$ with brown eyes, can be calculated based on Kolmogorov's probability axioms and the statistical principal principle ([sec. 4.7.2](#) – [4.7.3](#)). In summary, Popper's criticism of logical induction for empirical law hypotheses is unjustified or at least exaggerated; but it is certainly justified in the case of scientific theories.

Based on Popper's criticism of logical induction, we can evaluate competing theories only *comparatively* , that is, in relation to each other. Once, we have carried out an empirical success comparison of currently available alternative theories, we will base our future predictions and actions on those theories which up to now have been *most successful* , in other words, on the ones which have been corroborated best. Popper, of course, also sees it this way (1983, 65; 1979, sec. I.9). It follows from this that Popper's "deductivist" program of corroboration, too, contains a fundamental inductive step – a third kind of induction, that has been called epistemic induction ([Musgrave 2002](#)) or meta-induction ([Schurz 2008b](#)):

3 Epistemic induction or meta-induction says the following: If a theory T_1 has been more successful than T_2 *so far* (explanatory and prognostically), then it is probable, *relative* to the given state of evidence, that T_1 will also be more successful than T_2 *in future* . In other words, the success preferences established *so far* are projected inductively into the future. This induction principle is called "epistemic" or "meta-," because it is not object-level hypotheses (about ordinary events), but epistemic meta-hypotheses about the degree of corroboration of our object-hypotheses, which are being inductively projected. The epistemic induction principle is indeed indispensable for all empirical disciplines. Were this principle not to be accepted, there wouldn't be any point in Popper's method of testing. Success so far would then simply be irrelevant to our future actions. Although for example the view that heavy bodies on Earth are disposed to fall to the ground and not to float freely has been better corroborated so far than the opposite view, this would not be a reason to make this view the basis of our future actions, since logically, all bodies might just as well start floating as of tomorrow. Unfortunately, Popper kept an exaggerated distance to logical empiricism throughout his life,

and condemned any form of induction whatsoever. But his arguments on this topic in (1983, e.g. 66f) seem to rest more on mere terminological maneuvers than on actual material reasons (see [Schurz 2002a](#)). Other critical rationalists like Watkins ([1984](#) , 340f) and Musgrave ([2002](#)) have accepted the epistemic induction principle.

Epistemic induction, let us remember, is always only *comparative* . An important consequence of this is the following: according to epistemic induction, the degree of confirmation of a theory is always *doubly* relative:

- relative to the given state of observational knowledge, and
- relative to the current state of alternative theories.

2.7.3 Abduction and Inference to the Best Explanation

Scientific theories differ from empirical law hypotheses in that they contain theoretical concepts which speak about unobservable entities or properties. Inductive generalization enables us to infer from particular observations to empirical law hypotheses, but not from empirical law hypotheses to scientific theories. This is because inductive inferences cannot introduce new (theoretical) concepts into the conclusion, which are not contained in the premises. There is, however, one further – but controversial – kind of inference, by which this is possible, namely abduction, or inference to the best explanation. Roughly expressed, this is an inference from an observed *effect* to a likely *cause* , for example, from a curved trail in the sand to the recent passage of a sand viper.

At this point, we must warn against a *terminological misunderstanding* . A number of authors understand the concept of induction in a *broad* sense – they identify inductive inferences with *any* kind of content-expanding or nondeductive inferences, and so regard abduction as a form of induction (e.g. [Pollock 1986](#) , 42; [Bird 1998](#) , 13; [Ladyman 1998](#) , 28). By contrast, for the sake of a clear demarcation between distinct concepts, we understand induction always in the above outlined *Humean* sense, and see abduction as a second nondeductive or content-expanding kind of inference.

The abductive form of inference goes back to C.S. Peirce. In (1878a), Peirce explicated abductive inference formally as a retrodictive inference from “all Fs are Gs” and “this is a G” to “this is an F.” While this characterization of abduction does dominate the computer science literature on abduction (see [Flach/Kakas 2000](#)), it is much too narrow to grasp what the late Peirce was trying to achieve. The essential thing for Peirce ([1903](#) , §170) was that abductive inference allows new theoretical concepts and models to be introduced in the sciences. Newton, for example, inferred abductively from the movement of the planets around the Sun to the existence of a gravitational force. The epistemological inference from subject-internal perceptual experience to external empirical facts is also an abductive inference: the assumption that my visual experiences are being caused by objects existing independently of me is the best explanation for the constantly recurring regularities in my visual experiences (cf. [Schurz 2024 a](#), ch. 9). As is shown in Schurz ([2008a](#)), abduction comprises a whole *family* of kinds of inferences, which have the following pattern in common (cf. [Peirce 1903](#) , §189; [Niiniluoto 1999](#) , S439):

(Prop. 2.7-1) *General pattern of abduction* :

Premise 1 : A (singular or general) fact E, in need of explanation.

“ *Premise* ” 2 : A background knowledge K, which implies for a hypothesis H, that H is a potential and sufficiently plausible explanation for E.

Abductive conjecture : H is true.

As Peirce himself emphasized, the trustworthiness of an abductively acquired hypothesis is rather preliminary: the abducted hypothesis must go on to be empirically tested by deduction and induction in order to acquire the character of a probable hypothesis (1903, §171). More importantly, the given background knowledge always contains *several* possible hypotheses, H_1 , ..., H_n , which *potentially* explain the given evidence E (i.e. if they were true, they would explain E); and the abductive inference selects the most

plausible hypothesis among them. In this sense, Harman ([1965](#)) transformed Peirce's concept of abduction into the following inference to the best or best available explanation (IBE):

(Prop. 2.7-2) *Inference to the best (available) explanation – IBE* : Infer abductively to the most plausible explanatory hypothesis among all explanatory hypotheses for the given explanandum E that are possible, or are available in the given background knowledge, respectively.

Since the later Peirce also regarded abduction as IBE (1931ff, CP 6.525), abduction and IBE can be seen as approximately equivalent forms of inference (cf. [Douven 2011](#) ; [Williamson 2016](#) ; [Niiniluoto 2018](#) , 17f); however, this equivalence thesis is controversial (cf. the discussion in [Schurz 2023c](#) , ch. 1).

According to the understanding of Harman's ([1965](#)) and other representatives of IBE ([Lipton 1991](#) ; [Armstrong 1983](#)), an IBE infers from the explanatory success of a theory to its *likely truth* in the realist sense. Lipton ([1991](#) , 58) and van Fraassen (1989, 142ff) have pointed out that one can infer only to the best *available* explanation, because no one knows all possible explanations. In this situation, Kuipers ([2000](#) , sec. 7.5.3) has made an interesting suggestion. He brings the same objection into play as Popper made against logical induction: De facto, we can only ever reach *comparative* judgments, according to which a theory is explanatorily more successful and, hence, likely to be closer to the truth than all alternative theories *known* to us. Since this explanatory success includes also predictive success, Kuipers speaks more generally of "empirical success." Kuipers therefore suggests modifying the pattern of abduction in the following way: first, instead of speaking of an "inference to the best explanation," he speaks of an "inference to the best theory," and second, that a theory T is empirically most adequate among all known alternative theories does not lead to the conclusion that T is likely to be true, but only that T is *closest to the truth* among all known alternative theories. The interplay of the epistemic inductive inference and the abductive inference to the best theory is shown in [fig. 2.7-2](#) .

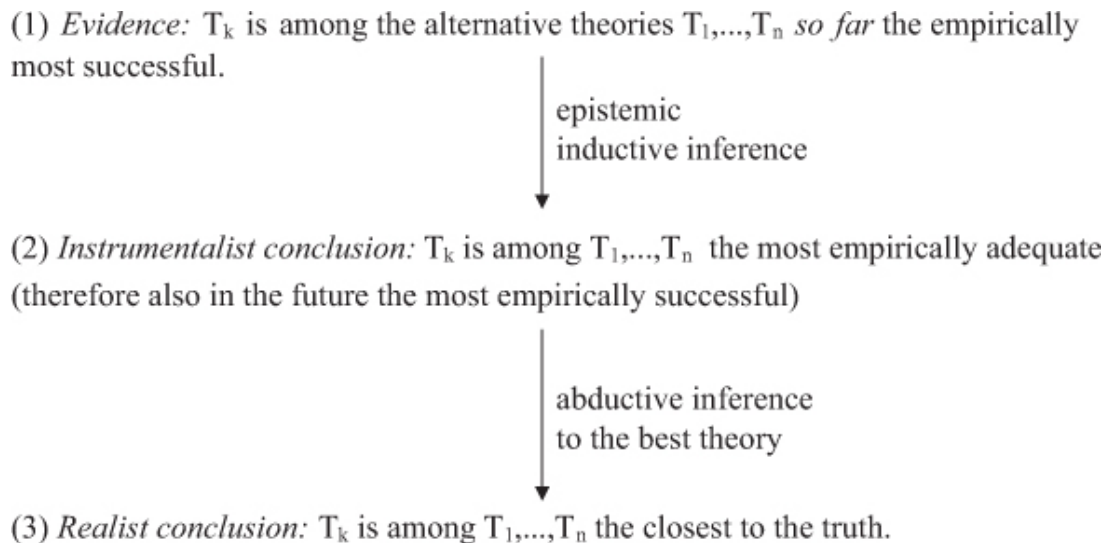


Figure 2.7-2 Interaction of epistemic induction and abduction. ↵

In [fig. 2.7-2](#) it becomes clear how these two kinds of inference are connected with the positions of *instrumentalism* and *realism* in the philosophy of science (see [sec. 5.7](#)). Instrumentalists, like for example van Fraassen, accept the inference from (1) to (2), but reject the abductive inference from (2) to (3), because according to them, theories may only be justified to be more or less empirically adequate, but not to be true or false in the realist sense (van Fraassen 1980, 11f, 20; 1989, 142f). Realists in the philosophy of science believe on the other hand in the likely truth or truthlikeness of well-confirmed theories; for them, it is that very inference from (2) to (3) which is essential.

2.7.4 Monotonic and Non-monotonic Inferences

Because deductive inferences are certain, that is, preserve truth (from premises to conclusion) in all possible worlds, they are monotonic in the following sense:

(Def. 2.7-1) *Monotonicity* : A valid inference (from premises $P_1, \dots, P_n$ to a conclusion C) is called monotonic, iff it remains valid after the addition of arbitrary further premises.

By contrast, inductive or abductive inferences don't possess the monotonicity property, because they are uncertain, that is, they don't transfer truth in all, but only in sufficiently "normal" worlds. We restrict the notion of *validity* to deductive inferences (indicated by a single forward slash "/"), while in the case of uncertain inferences, we speak instead of their (nondeductive) *correctness* (indicated by a double forward slash "//"). With this notation, monotonicity means that if " $P_1, \dots, P_n / C$ " is valid, then " $P_1, \dots, P_n, Q / C$ " is valid for every new premise Q .

The non-monotonicity of uncertain inferences makes it possible that " $P_1, \dots, P_n // C$ " is correct, but " $Q, P_1, \dots, P_n // C$ " is incorrect for some (true) premise Q , where Q typically expresses "exceptional information." For example, "1000 swans have been observed to be white // Therefore probably: All swans are white" is a correct inductive generalization, but the premise-expanded inference "1000 swans have been observed to be white and 10 to be black // Therefore probably: All swans are white" is inductively incorrect. The logical examination of *non-monotonic* inferences has become a significant area of inquiry since the 1980s (cf. [Gabbay et al. 1994](#) ; [Schurz 2004](#)).

A non-monotonic inference with true premises has only *preliminary* value and can be made worthless by new (true) information. In epistemology, non-monotonic inferences are therefore also referred to as *defeasible* (cf. [Pollock 1974](#) , 36ff). The preliminary nature of non-monotonic inferences is expressed by the fact that they have a different *epistemic application rule* than monotonic inferences. Here, we take the *belief system* (or 'knowledge' system) S of a person or group of persons to be the set of propositions which are justifiably accepted as true by these persons.

(Prop. 2.7-3) *Application rule for monotonic inferences* : If a deductive inference is valid, then you should add its conclusion to your belief system S , as long as S contains the premises.

(Prop. 2.7-4) *Application rule for non-monotonic inferences* : If an uncertain inference is correct, then you should add its conclusion to your belief system S , as long as S contains the premises, and these premises cover *all* propositions in S which are of relevance to the conclusion. For inductive-probabilistic inferences, [prop. 2.7-4](#) is equivalent to Reichenbach's rule of the narrowest reference class and Carnap's requirement of total evidence (see [sec. 3.9](#) , [3.13.4](#) , [6.3.2](#)).

Complementary and Advanced Topics

2.8 How Much Metaphysics is Presupposed by Science?

2.8.1 Truth, Idealism, and Realism

One may object that our reconstruction of the supreme goal (G) of the sciences and the assumption ($E1$) of minimal realism assumes too much of realism. Famous scientists and philosophers of science, such as George Berkeley (1685–1753), Ernst Mach (1836–1916), and the early Rudolf Carnap, were idealists or positivists, that is, anti-realists. In their view, science is possible if it is merely based on human intersubjective experiences, *without* the presupposition of a subject-independent reality that transcends these experiences. We do not deny this possibility.

Admittedly, our assumption ($E1$) makes a (weak) realistic assumption, which is discussed below. Some philosophers argue that already our explication of the goal (G) of science, the discovery of content-rich

truths, presupposes realism, insofar as truth means correspondence of statements with an assumed reality. But this is not the case, since (G) does not say that the domain of phenomena to which true statements correspond has to belong to an external and subject-independent reality, as claimed by realists. The domain of phenomena called ‘reality’ may also consist of experiences, as assumed by idealists.

The major epistemological positions can be summarized as follows (see [Dancy 1985](#) , 136ff; cf. [Schurz 2024a](#) , [sec. 2.7](#)). The domain of entities to which true statements correspond consists either of

- a. one’s own subjective experiences (*subjective idealism* or solipsism),
- b. the experiences of all subjects (*intersubjective idealism*),
- c. the *possible* experiences of all subjects (*possibilistic idealism*), or
- d. a subject-independent reality that causes the experiences (*realism*).

Solipsism is a rather extreme position (most idealists have been possibilistic idealists). Be that as it may, the important point to observe here is that the concept of *truth* makes sense in *all* of these positions. Hence, truth in the sense of correspondence theory is not realistically biased but *metaphysically neutral* (cf. [Kirkham 1992](#) , [sec. 6.5](#)). So also, our formulation of the supreme goal (G) of science is neutral. It does not entail that successful science presupposes the epistemology of realism. The goal of truth makes sense for *all* of the epistemological positions; only the *kinds* of truths that, according to these positions, are pursued by science are different (true descriptions of experiences versus true descriptions of subject-independent facts).

While (G) is metaphysically neutral, our epistemic assumption (E1) is not. It asserts that realism, in the explained weak version of it, is the *normal* epistemological assumption of science. I think this claim is defensible: more or less all empirical scientists assume the external reality of their objects of investigation; idealist or anti-realist worldviews are the clear exception in normal science, although the realistic attitude of normal science can be rattled in scientific revolutions (recall [sec. 2.1](#)).

What distinguishes our minimal realism from all sorts of *metaphysical realism* is that neither the existence of a subject-independent reality, nor its knowability, is assumed to hold on a priori reasons or is a necessary presupposition of knowledge. By contrast, the minimal realism of assumption E1 is a *hypothetical-constructive* realism, insofar as we consider realism not as a presupposition, but as a hypothesis or model construction – in analogy to scientific hypotheses. Our realism is committed to *methodological naturalism* , as the hypothesis of realism must be corroborated by the *empirical success* of our cognitive belief system, as the *best explanation* of this success. This abductive form of realism corresponds at the same time to a variant of empiricism, namely *abductive empiricism* , which – unlike the reductionist empiricism of the neopositivists – is compatible with realism ([Schurz 2024a](#) , 90).

Because of its naturalistic tenets, hypothetical-constructive realism is compatible with legitimate doubts about a ‘full’ realism found in the philosophical interpretations of *quantum mechanics* . In all of its interpretations, quantum mechanics assumes the existence of a microphysical reality that possesses some subject-independent and discoverable properties (so-called observables). Only certain properties of entangled (or superposition) states are not objectively measurable, because it is a postulate of quantum mechanics that every measurement transforms an entangled state into a statistical ensemble of pure states (see [Herbert 1985](#) , sec. 8+9; [Penrose 2007](#) , sec. 29).

Hypothetical-constructive realism is also compatible with *structural* realism. In its *epistemic* version, structural realism says that only the structure but not the intrinsic nature of the entities making up reality are knowable ([Worrall 1989](#) /1997; [Psillos 1995](#) ; [Votsis 2011](#)). The *ontic* version of structural realism is motivated by quantum mechanics and claims that even the entities that make up reality are themselves not individuals but structures ([Ladyman 1998](#) ; [French/Ladyman 2003](#)).

2.8.2 Constructive Realism versus Radical Constructivism

Hypothetical-constructive realism differs markedly from radical *constructivism* mentioned in [sec. 1.3.9](#) . The latter position was developed by psychologists such as Glasersfeld ([1995](#)) or biologists like Maturana and Varela ([1992](#)). Historically speaking, radical constructivism is a modern variant of subjective idealism. The main argumentation of this philosophical position is based on the fact that our perceptions and concepts of reality are cognitive *constructions* of our mind or brain, respectively. Radical constructivists, however, infer from this hardly deniable premise that even ‘reality’ itself is a construction, that is, a reality that is

independent of our constructions does not exist. A similar argument has been put forward by social constructivists such as Latour and Woolgar ([1986](#)), in regard to the ‘construction’ of scientific facts by scientists.

In this subsection, we want to show that the argument of radical constructivists is based on a *fallacy*. We reconstruct this argument as follows (e.g. [Glaserfeld 1995](#), 40, 51; [Kriz et al. 1990](#), 19f):

(2.8-1) *Premise of radical constructivism = epistemic constructivism* : Our perception and mental representation of reality is not something given, but the result of an active cognitive construction.

Conclusion of radical constructivism = ontological constructivism (idealism) : Therefore reality, too, is not given “in itself,” but “constructed” or “constituted” by us.

The crucial fallacy of this argument is the ‘magical transformation’ of epistemic into ontological constructivism (cf. [Sokal/Bricmont 1998](#), 102). Hypothetical-constructive realism shares the premise of ([2.8-1](#)) with radical constructivism, but it definitively rejects the conclusion. The *invalid* step from epistemic to ontological constructivism is frequently substantiated by constructivists with the help of the following ‘justification’ (cf. [Kriz et al. 1990](#), 20; [Glaserfeld 1995](#), 51), which may be called the *Kantian fallacy*, because it corresponds to a popular interpretation of Kant’s philosophy:

(2.8-2) “ *Justification* ” of radical constructivism : It is impossible to know something about reality *in itself* – independent of its epistemological construction. The only thing about which we can know something is the reality *constructed by us*.

The subtle error of this justification attempt is that the concept of “knowing” is understood in the *naïve-realist* sense of a kind of direct mirror-imaging. Then it is indeed impossible to know something about an object that is “independent of our epistemological construction” – which in this case means: “ *not influenced* by this construction.” Modern truth theory, though, understands the relationship between a true statement and its object not as a quasi-identical mirror-image, but as a *structural correspondence*, which *conveys certain information* and has to be neither complete nor unique. In this view, there is no incoherence in saying that *by means of* our epistemological constructions we know something about a reality which exists *independently* of our epistemological constructions. For the independently existing reality is not taken to be immediately given to our mind, as in metaphysical or ‘direct’ realism, but is postulated as a system of hypothetical entities, whose existence can best explain the factual success of our knowledge. Giere ([1999](#), 174ff) and Kuipers ([2000](#), sec. 13) have called this position “constructive realism,” and we call it *hypothetico-constructive* realism or alternatively *abductive* realism, in order to emphasize that realism too is ultimately a fallible hypothesis.

In conclusion, it is the very argumentation of radical constructivists which clings to naïve-realism, inasmuch as it uses a naïve-realist interpretation of the objects of our knowledge to prove that these objects are subjectively constructed; while from the hypothetical-realist view, the statement ([2.8-2](#)) and so the whole argumentation of radical constructivists is untenable.

2.9 On the Theory (In)dependence of Observation

Classical empiricists argued that all scientific (or cognitively valuable) concepts are reducible to observational concepts. That this is not so was convincingly demonstrated by the recognition of *theoretical concepts* in the middle and later stages of logical empiricism (cf. [Carnap 1936](#) /37, 1956a). Theoretical concepts go beyond what is observable – this is a generally accepted insight in post-positivistic philosophy of science ([sec. 5.1](#)). Accordingly, the empiricism in our reconstruction of the common epistemological ground of science ([sec. 2.2](#)) is reduced to a minimalistic version according to which theories, although going beyond what is observable, must at least have *some* observable consequences by which they are testable.

What minimalistic empiricism still assumes is the existence of at least *some* theory-independent observation concepts and sentences, which can be used to independently test competing theories. However, even this minimalistic empiricism came under attack in the post-positivistic philosophy of science. The thesis of the *theory-dependence* of all observation and experience has been made popular by Hanson ([1958](#)), Kuhn ([1962](#)), and Feyerabend ([1975](#)); with precursors in Wittgenstein ([1945](#) /2001, part 2, §xi) and Popper (

[1935](#) /2002, new appendix X). Meanwhile, this thesis seems to be accepted by the majority of philosophers of science. Closer inspection reveals, however, that what is accepted by the majority of philosophers is a cautious and loose version of this thesis, that leaves plenty of room for interpretations that are compatible with our epistemological assumptions (E3) of objectivity and intersubjectivity, and (E4) of minimal empiricism. For many philosophers of science, what is important in the theory-ladenness thesis is a critical attitude toward what is called “data” in science: these data are typically theory-laden interpretations of the scientist’s perceptions, and the assumptions made in these interpretations are often concealed behind a veil of expertocracy. Nevertheless, I will try to show in this section that in all empirical sciences, one can find theory-neutral observation concepts and statements that scientists may draw on in cases of conflicts of opinion.

Before I start my enterprise, let me try to demonstrate the significance of this question for the philosophy of science. A *complete* theory-dependence of observations, in the sense that *all* observations are theory-dependent, would turn empirical science into a *circular* affair without any hope of objectivity (cf. [Nagel 1979](#) , 79). For suppose that two adherents of rival theories T_1 and T_2 are observing the same area of space-time, and assume that their respective theories imply the contradictory predictions P_1 and P_2 . If complete theory-dependence were given, then the observations of the theory-adherents would be determined by their theoretical *expectations*. The supporter of T_1 would then observe P_1 where the supporter of T_2 observes P_2 . So, each prediction would be a *self-fulfilling* prediction, and an independent empirical test of theories would be impossible.

Complete justificatory circles are demonstrably epistemically worthless (see [sec. 2.11](#)). By contrast, a *partial* theory-dependence, in the sense that there are at least some theory-neutral observations on which theory-dependent scientific data rest, is not a threat to the possibility of objective empirical tests. Adam ([2004](#)) and Brown ([1993](#)) have studied cases in which the recognition of a scientific datum D that falsifies a theory T relies on the assumption that theory T is true. This sounds paradoxical, but in fact is not: the case is harmless, because the datum D was gained from a theory-neutral observation O using T . This means (in the simplest case) that T entails the implication *if O , then D* . Moreover, that T is falsified by D means that D entails not- T . Taken together, this implies by simple logical means that O entails not- T , that is, that the observation O falsifies T . Nothing paradoxical is involved in this situation.

In conclusion, for a defense of the objectivity of empirical science, it is sufficient to refute the thesis of complete theory-dependence, that is, to establish the existence of *some* theory-neutral observation concepts and statements in science. In the next subsection, I will sketch eight central arguments in favor of different forms of theory-dependence as well as the corresponding counterarguments. I will try to show that although the pro arguments make correct points, they do not exclude the existence of theory-neutral observations in the narrow sense of perceptions.

2.9.1 Theory-Dependency of Observation – Pros and Cons

Pro-1 : Observations are expressed as realist statements . Observation statements such as “this thing is green” are usually expressed in the realistic mode. So, they imply the *hypothesis of realism* , that is, the existence of subject-independently existing objects (“this thing”) that cause our subjective perception (cf. [Popper 1935](#) /2002, sec. V, §25, for a version of this argument). Observation sentences in this realist interpretation transcend the content of subjective experience. We can be in *error* about them, for example, because our sense experience may rest on a *hallucination* .

Contra-1 : This argument is true, but this is not what is meant with theory-dependency of observation. What is meant is that the *content* of one’s observations depends on the *content* of one’s background theory (expectation or worldview). But the question whether one is an epistemological realist or a phenomenalist idealist does not affect the content of one’s observations. It merely changes the interpretation of the term “thing:” in the realist interpretation, this term refers to assumed entities outside of one’s experiences, and in the phenomenalist interpretation, it refers to entities within one’s experiences. But whether one is a realist or an idealist does not turn a green-observation into a blue-observation (this argument can be found in [Carnap 1928b](#)).

Pro-2 : Observations are theory-guided . The selection of relevant problems and experiments is essential in every scientific research, for reasons of complexity alone. This selection is guided (among other things) by one's theory: which theory I support determines which observations I look for. ³

Contra-2 : It is true that the selection of one's observations depends on one's theories, and more generally, on one's interests. But that observations are selected in a theory-dependent way does not imply that the observations *themselves* are theory-dependent, in the sense that supporters of different theories would observe different things when perceiving the same spatiotemporal region or measurement result. Supporters of the theory T_2 *could* also make the observations made by supporters of the theory T_1 and *would* then recognize their truth. More generally speaking, the fact that empirical research is interest-driven does not imply that certain observations *cannot* be made, but at most that these observations are not *wanted* to be made, or that they are *ignored* . However, such a dogmatism is a sign of *bad* science and violates the principle of *total evidence* ([sec. 3.13.4](#)).

Pro-3 : Perceptions are unconscious theoretical constructions. That perceptions or observations in the *narrow* sense are unconscious theory-like constructions is an insight that goes back to Helmholtz ([1896/1910](#)). The brain's visual regions use refined neural construction mechanisms to create a consciously experienced 3D image from the unconscious 2D retinal light-impressions. Normally, this construction process generates a unique 3D image within milliseconds, so that its constructive character is *hidden* from the consciousness: you think you see "reality itself" (for this reason, humans are intuitively naïve-realists). But in certain – typically in artificially engineered – situations, this visual construction process reveals its constructive character, because it leads to consciously experienced visual *ambiguities* , *illusions* , or other strange effects. This has been impressively demonstrated by the findings of cognitive psychology (cf. [Rock 1984](#)). Simple examples of ambiguities are *tilting pictures* such as Jastrow's duck-rabbit, a picture that can be seen either as a duck or a rabbit (see [fig. 2.9-1](#)). The duck-rabbit has been used by Wittgenstein ([1945/2001](#), part II, §xi), Hanson ([1958](#) , 12), and Kuhn ([1962](#) , 112ff) as evidence for the 'theory-dependence' of perception, insofar as what we *expect* to see influences what we see.



Figure 2.9-1 Duck-rabbit. [↗](#)



Figure 2.9-2 Mueller-Lyer illusion [↗](#)

An example of an *optical illusion* is the Mueller-Lyer illusion in [fig. 2.9-2](#) . The explanation of this illusion lies in mechanisms of 3D-vision: the arrows and inverted arrows produce the 3D-effect that the right stroke appears to be closer to the observer than the left stroke and is therefore perceived as being shorter, although both strokes have equal length (see [Rock 1984](#) , 163, 167).

Contra-3 : Neither the logical empiricists nor Popper had good replies to the psychological arguments concerning the constructive character of perception, because according to their position of "logical purity," philosophy should be independent from psychology. What counts as an observation sentence was logically speaking a matter of "pure convention" (cf. [Popper 1935/2002](#), sec. 5, §30; [Neurath 1934](#)). We now present some arguments against the theory-ladenness of perceptions that are themselves based on findings from cognitive psychology.

The findings about the constructive character of visual processes refute only *direct realism* ([Dancy 1985](#) , 144). According to this epistemological position, we see things directly, as they are 'in themselves,' without

any mediating processes that construct intermediate images. In the light of contemporary knowledge about the neuro-cognitive process of vision, direct realism seems to be an untenable position (see [Schurz 2024a](#) , [sec. 3.5](#)).

However, all findings of cognitive psychology also confirm that our visual perception processes and their results are quite stubbornly *independent* of our *learned* information. Everyone experiences the right stroke in [fig. 2.9-2](#) as shorter than the left one, even if she has been informed that this visual experience is an illusion. These facts motivated Fodor's thesis of the modularity and impenetrability of vision ([Fodor 1984](#) , 34ff). Also, the Helmholtzian constructivist Rock ([1984](#) , 228) acknowledges the independence of vision from acquired information.

In the tilting picture of [fig. 2.9-1](#) , everybody is able to perceive it either as a duck or as rabbit and to switch back and forth between these two gestalts. Kuhn argued that a similar gestalt switch took place when astronomers saw Uranus before 1781 as a fixed star, but after 1781 as a moving planet (1962, 111, 121f). But it does not seem right to conceive the Uranus case as perceptual switch: The visual content of the Uranus-perception is always a bright spot in the sky, which is perceptively stable and does not switch. The question whether it represents a fixed star or a planet is not a question of perception but of theoretical interpretation.

The processes of human vision are based on innate mechanisms that have developed in millions of years of evolutionary history. Their selection history explains why the *same* mechanisms, which under normal environmental conditions give a correct visual depiction of reality, can create optical illusions under certain non-normal conditions. Evolutionary considerations also explain why human perception, in order to be evolutionary successful, must be independent from expectations, for otherwise it could not correct *wrong* expectations. As an example, consider prehistoric men entering a cave which is inhabited by a cave bear that they did not expect: if their perception would not immediately have corrected their wrong expectation, these prehistoric men would not have survived.

By "theory-dependence" of perception, one could at best understand that perception is dependent not on learned information, but on 'innate theories' which all humans share, because they are rigidly wired up in their visual cortex. However, we do not want to understand the concept of "theory" in such an overstretched way. By the theory-dependence of perception, we always only mean the dependence of perception on *acquired* information that may vary in different subjects, cultures, or scientific communities. It seems that visual perception in this sense is – at least to a large extent – theory-independent.

Pro-4 : Sometimes, visual perception depends on one's pre-expectations . The last statement of Contra-3, that visual perception is independent from pre-expectation, is not generally true – it holds only for clearly presented visual stimuli. There exist some psychologically well-investigated experiments, to be discussed below, in which visual perception does indeed depend on pre-expectations.

Contra-4 : This is only the case for the stimuli at the perceptibility threshold , that is, when the limits of perceptual recognition capacities are reached (cf. [Brewer/Lambert 2001](#) , S181). These limits are reached in three kinds of cases:

1. *Poorness* of the stimulus: For example, (1.1) in a dark forest at night, anxious persons seem to see the contours of mysterious figures in the dark thicket, which disappear at daylight. Or (1.2), if the relevant stimulus is too small to be seen, for example, when you ask which of two approximately equally long strokes is the longer one, the test persons' answers may become expectation-dependent. The perceptibility threshold is also reached when (1.3) the presentation time of the stimulus is very short, as in high-speed reading tasks (cf. [Estany 2001](#) , 208).
2. *Ambiguity* of a stimulus that has no unique gestalt solution, such as the duck-rabbit in [fig. 2.9-1](#) : Here, one's expectation about the gestalt influences which gestalt one sees first (though one may switch the gestalt at will).
3. A *high complexity* of the relevant stimuli patterns is a third reason, not for expectation-dependence, but for dependence on acquired knowledge: here, visual *training* may be necessary to detect the relevant visual patterns. As an example, think of the radiologist who recognizes pneumonia in the X-ray image of your chest.

Because of these finding, some cognitive psychologists described human vision as a combination of bottom-up processes (information flow from the retinal cells) and top-down processes (information flow from higher

cognitive regions; cf. [Estany 2001](#) , sec. 4). Although this description is not false, it misses the important point of this combination, namely the *priority ordering* in favor of bottom-up information. Top-down information from expectation by memory can correct visual bottom-up information *only* if this bottom-up information is *too poor* to allow a clear gestalt interpretation. Otherwise, bottom-up information *overrides* top-down expectations. This is illustrated, for example, by an experiment in which a test person looks through a peephole with one eye at two white balls of equal size and distance in an otherwise dark black box ([Rock 1984](#) , 80f). In this experiment, all cues about the distance of the two balls coming from bottom-up information have been eliminated: there is neither (a) binocular depth information nor (b) monocular depth vision (e.g. by motion parallax via eye movement). If one tells the test person that one is a tennis ball and the other a ping pong ball, the tennis ball is perceived as being further away. This illustrates the utilization of acquired information when bottom-up information is missing. However, as soon as cues of sort (a) or (b) are admitted, for example, when the peephole is broadened, and the person is allowed to move her eye or to look with two eyes, the illusion disappears, and the person sees the two balls in equal size and distance.

Pro-5 : Scientific data (or observations i.w.s.) are theory-dependent. Scientists commonly use a wider concept of ‘observation’ or ‘datum’ than perception with the ‘naked eye.’ According to the scientific usage of this word, scientists observe galaxies, bacteria, or electrons with suitable *instruments*. A number of authors, including Kuhn in many passages, relate their thesis of theory-dependence to this observation concept i.w.s. [Giare \(1999 , 12f, 80\)](#) and [Kuipers \(2000 , sec. 13.3\)](#) argued that whenever a new measurement instrument is invented, a formerly theoretical concept is turned into an observational concept.

Contra-5 : Scientific data are theory-dependent interpretations of perceptions. The theories presupposed here are mainly theories about the way measurement instruments work, which are usually treated as an unproblematic part of the advanced knowledge of a scientist. This *pragmatically* extended observation concept i.w.s. is undoubtedly very important for the efficient progress of science, since one cannot always put everything simultaneously into question, and the investigation of scientific problems rests on large amounts of background knowledge that is highly corroborated and taken for granted. However, this fact does not imply what some philosophers of science wrongly infer from it, namely that there *are* no observations in the narrow sense (i.n.s.) in science, that is, direct *perceptions* (mainly visual perceptions). We claim theory-independence only for this narrow concept of observation.

Most of the time, scientists agree on their observations i.w.s., and there is no need for them to make their direct perceptions explicit. But when scientists disagree on what they have observed i.w.s., they will critically reflect their implicit theoretical assumptions that led them to infer different observations i.w.s. from their common perceptions. For example, a pulmonologist A may report to have seen pneumonia in the X-ray image of a patient’s chest. What he has seen directly is a certain blackening pattern that he interprets as pneumonia. If another pulmonologist B doubts this diagnosis, then what B doubts are not the visual abilities of A. Rather, A and B will discuss whether the blackening pattern which they *both* see is indicative of pneumonia or not. In conclusion, the epistemological significance of observations i.n.s. lies not in being permanently verbally reflected, but in serving as a possible level of theory-neutral agreement to which scientists can step back in the case of disagreement about observations i.w.s.

Pro-6 : There is a continuity between observational and theoretical concepts. Maxwell ([1962](#) , 7), Carnap ([1966](#) , 255f), and Hempel ([1966](#) , 78–84) pointed out that there is a continuous transition from observability with the naked eye to intermediate stages such as observability with glasses or a magnifying glass, until observability with an optical telescope or microscope, an electron microscope, or a radar screen. A number of authors inferred from this that any demarcation between perception and theory-dependent observation i.w.s. is arbitrary.

Contra-6 : That there exists a continuous transition in the extent to which concepts are theory-laden is true. But first , the fact that there exists a continuous transition between two features like for example black and white, does not by any means imply that there is no difference between black and white (cf. van Fraassen 1980, 214). So, the distinction between observation in the narrow and in the wide sense is significant even if it is continuous. *Second* , there are marked jumps in this transition: while observation i.n.s. does not require any learned background knowledge, the required background knowledge increases successively for the usage of the above-named instruments. As an illustration of this point, imagine what a native tribesman would do with these instruments, without having received instructions. Presumably, he would understand glasses, but

already optical microscopes and telescopes would hardly be understandable, and that's not to mention electronic microscopes or radar screens, as the tribesman has no clue about the background knowledge that is needed for an adequate handling of these instruments.

Pro-7 : Meaning holism makes observation concepts theory-dependent. Quine ([1951](#)) doubted the traditional analytic-synthetic distinction (see [sec. 3.11](#)). His alternative semantic view was a radical meaning holism, according to which the meaning of a concept is determined by all that we know about it. If this were true, then theory-neutral observation concepts could not exist, because in scientific theories, observation concepts get connected with theoretical concepts. According to meaning holism, this would imply that their meaning becomes loaded with theoretical statements. In this sense, Chalmers ([1986](#) , 36) argued that “water” is a theory-laden concept, because chemical theory connects it with the molecular structure H_2O .

Contra-7 : Radical meaning holism will be criticized in [sec. 3.11](#) as more or less *untenable* . Quine's inner conflict concerning observation concepts is revealed in (1992, 7f), where he attributes a “two-faced nature” to the sentence “this is water:” “holophrastically,” so Quine, the sentence is free of theory, but by its composite structure, the sentence contains theory. This sounds like an ad hoc maneuver to escape the untenable consequences of meaning holism. We argue that the meaning of observational concepts is fixed by processes of definition or ostension such as “this is water.” This meaning is *not* affected by theories, which connect observational with theoretical concepts. If this were not so, then one would have to say that a physicist and a layperson mean entirely different things when they point to a creek and say that water is bubbling there – but this seems too strange to be true.

Pro-8 : Observation concepts (as all other concepts) are language-dependent. The thesis of the language dependence of experience was made popular by the philosopher Wilhelm von Humboldt and by the anthropological and ethnological research of Sapir and Whorf (see [Kutschera 1976](#) , [sec. 4.1](#)). Many native tribes, for example, have different and less differentiated concepts of color than members of Western civilization ([Berlin/Kay 1999](#)). From this fact, cultural anthropologists inferred that these natives really see the colors in their environment differently from us. The other way around, Inuits (Eskimos) have reportedly many more concepts for sorts of snow than us (see [Whorf 1956](#) , 216; [Kutschera 1976](#) , 251f). Whorf derived from this his *linguistic relativity principle* , according to which we can only perceive what is describable, and in a sense also prescribed, by the concepts in our language.

Contra-8 : It must be admitted that different cultures, to a certain degree, develop different systems of concepts, in which they describe the perceptions of their different environments. But from this, it does not follow that their sense experience itself is language-dependent. We could only conclude this if members of these cultures were also not able to *learn* the observation concepts of other cultures by *ostensive training* , that is, by acts of pure ostension. This ostensive learning ability is present in all cases, however. The Zuni Indians can learn the concept of “orange,” for which they don't possess a word, through brief ostensive training – that is, through simple acts of pointing (“this is orange,” “this is not orange”). The Tiv and the Dani were able to learn many European color concepts in this way, for which there is no word in their mother tongue ([Garnham/Oakhill 1994](#) , 49–51; [Berlin/Kay 1999](#) , 24). Similarly, a European can learn the Inuit's words for different kinds of snow ostensively.

2.9.2 A Definition of Theory-Neutral Observation Concepts

In this section, we present our definition of theory-neutral observation concepts. We start from the insight that the defining criterion cannot come from logic, but must be an empirical-psychological one. Hence, the definition of “observation concept” will be relative to the sensorial capacities of humans. For example, the magnetic field is non-observable for humans, but observable for animals with magnetic sense organs (cf. [Carnap 1936](#) /1937, 454f; van Fraassen 1980, 56f).

Traditionally, an observational concept is characterized as a concept whose applicability to an observable object can be determined through mere perception (e.g. [Carnap 1936](#) /1937, 454f). To verify this characterization, we must rely on the *introspection* of individuals – but it is well known that many people indiscriminately conflate perception with theoretical interpretation in their supposed observations. For example, esotericists claim to observe “spiritual fields,” and 18th-century chemists reported observing

phlogiston. To avoid this problem, we base our definition on an experimentally testable criterion: the *ostensive learnability* of concepts. With this step, we do not only circumvent the problem of the unreliability of introspection, but also the influence of language. An *ostensive learning experiment* is based on nonverbal behavior and involves only three communicative acts that are culturally universal: acts of ostension (“this”), interrogative forms (“is this X?”), and words or gestures for “yes” and “no.”

It is essential for an *ostensive learning experiment* to work in a *nonverbal* way. Otherwise, the linguistic relativist could argue that the Zuni have learned to see the color “orange” only by learning the Western color language. Also, Quine says in several places that the characteristic property of observation sentences is their ostensive learnability, independently of acquired information (1960, 42; 1974, 37, 41f; 1992, 3f, 5f). But this thesis of Quine is in strong tension with Quine’s radical meaning holism (cf. pro-and-contra 8 of [sec. 2.9.1](#)). So, in spite of these similarities, Quine’s theory of observation sentences is essentially different from our account.

An ostensive learning experiment consists of a *training phase* and a *test phase*. In the training phase, the concept to be tested for observability is introduced to the test person (tp) with an *artificial word* “X” unknown to them, so that the tp does not associate *any* meaning with the artificial word. The concept referred to by “X” is illustrated to the tp with a set of positive and negative instances of the concept, and the information “this is an X” or “this is not an X” is given. The instances may be presented in the form of photos, movies, or real objects. They must be clearly visible by the tp; more precisely, the experiment takes place under ‘normal observation conditions’ (see below).

The ostensive learning process to be achieved by the tp consists in the abstraction of the meaning of “X,” that is, the property cluster corresponding to X, from the examples presented. Whether the tp has succeeded in this is investigated in the *test phase* which follows, in which a series of *new* positive and negative instances of the concept are presented, and the question “is this an X?” is asked. The tp’s success frequency in correct answers in the test phase indicates to which extent the tp has successfully abstracted from the training instances the perceptual property cluster that corresponds to the concept “X.”

If a concept can be learned from a small set of learning instances by (almost) all test persons, *independently* of their background culture and even if the tps possess no word for this concept in their spoken language(s), this is a more-or-less compelling reason to see this concept as a *theory-independent* observation concept. Without this assumption, the learning success would be a sheer miracle. If a Zuni Indian were unable to perceive the color feature “orange” because she has no concept of it, then she would not perceive the difference between the orange, red, and yellow pictures in the training phase, and it would be *inexplicable* how she could successfully pass the test phase. In conclusion, we propose the following definition of observation concept:

(Def. 2.9-1) A concept is a *theory-neutral observation concept* (or an observation concept i.n.s.) iff almost all humans can acquire this concept in an ostensive learning experiment, under normal observation conditions, independently of their background information, language, and culture.

Thereby, the *normal observation conditions* include (i) the normal physical conditions (sufficient light etc.), (ii) normal biological conditions (no biological deficiencies in the person’s sensory organs and nervous system), and (iii) normal psychological conditions (the person is in a wakeful and mentally normal state, and her yes/no utterances are truthful). All three conditions can be experimentally tested.

Our definition of observation concepts does not imply that all cultures have the same system of observation concepts (this is de facto false, as we know). What the definition implies is that different cultures can learn each other’s observational concepts in an ostensive manner, without being forced to adopt the language or worldview of the other culture. This consequence accords with the facts mentioned in [sec. 2.9.1](#) (contra-8), namely that native tribes can learn the Europeans’ color concepts, or that Europeans can learn the Inuit’s concepts for kinds of snow (etc.).

It would be interesting to perform ostensive learning experiments of the described sort (according to my knowledge, this has not been done so far). Generally speaking, the number of instances that are necessary to grasp the perceptual property cluster that corresponds to the concept “X” will not only depend on the question whether the property cluster is observable in principle or not, but also on the *complexity* of this property cluster. So, the result of an ostensive learning experiment is expected to be a *gradual learning curve*

. This turns the resulting notion of an observation concept into a *gradual* one: a concept may be more or less observational. Nevertheless, we expect that there will be clear statistical divide between the following types of concepts, as illustrated in [fig. 2.9-3](#) (cf. [Schurz 2015b](#)):

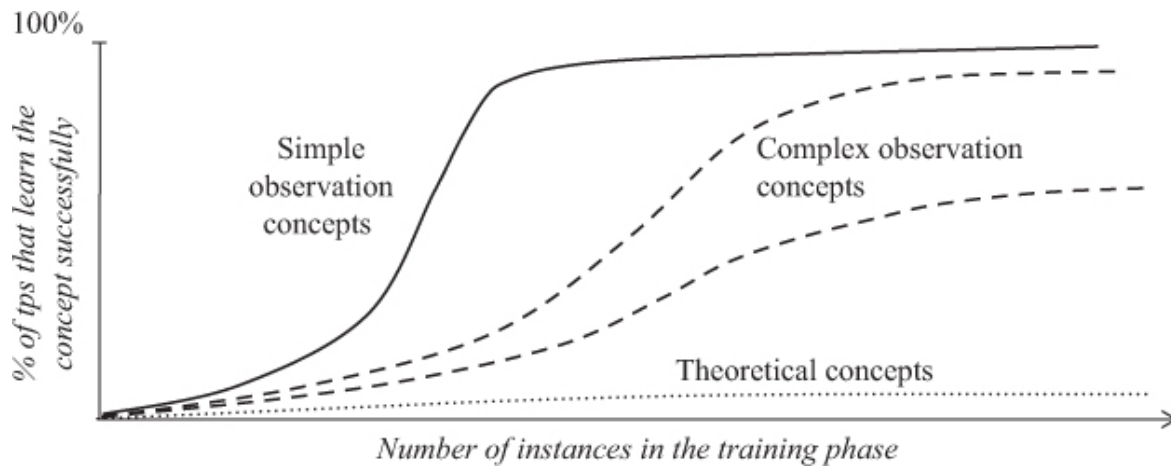


Figure 2.9-3 Expected learning curves in ostensive learning experiments. [↗](#)

1. For *simple* observation concepts (such as “red,” “square,” or “bird”), success frequencies will after a small number of instances climb above the significance threshold, toward values slightly smaller than 100% of tps.
2. For *complex* observation concepts (such as “either square and red, or circle and blue” or “hooded crow”), the success rate will climb slowly and will need many instances; often the success frequency will stagnate at lower levels, for example, at 50%.
3. For genuinely *theoretical* concepts (such as “atom,” “electric force,” or “oxidation”), the learning curve will not climb at all, even after many instances.

2.10 Value Neutrality in the Context of Contemporary Debates

2.10.1 The Role of Values in Science

As pointed out in [sec. 2.5.2](#), the value-neutrality requirement ([prop. 2.5-1](#)) possesses a critical-empiricist function. This stands in contrast to an influential division of epistemological paradigms according to their respective knowledge-guiding interests, which goes back to Habermas ([1966](#), 241ff) (cf., e.g., [Kriz et al. 1990](#), 151; [Rosenberg 2008](#), 127ff):

<i>Epistemological paradigm</i> :	<i>Knowledge interest:</i>
1 empirical-analytical	technical-instrumental controllability
2 hermeneutic	establishment of normative consensus
3 critical-dialectical paradigm	emancipation from power structures

Given our understanding of scientific value neutrality, these assignments are clearly incorrect. Science, being classified by Habermas as empirical-analytical, cannot be assigned to *any* specific interests that are external to science, but serves directly merely the science-internal epistemic goals, the search for content-rich truths. Scientific knowledge may be equally well used for technical purposes or for purposes of political reform. Adherence to value neutrality certainly does not mean that a scientist must refrain from practical involvement in society, or remain unaware of her social *responsibility* (cf. [Prim/Tilman 1979](#), 139f; [Koertge 2000](#), S48). All that the value-neutrality requirement implies is that the scientist *separates* her scientific knowledge from

fundamental value assumptions which she assumes in means-end inferences, and thereby *relativizes* her means-end inferences hypothetically to these value assumptions (recall the example in [sec. 2.5.2](#)).

The true difference between Habermas' three epistemological paradigms lies *not* in differing interests, but in the different treatment of the *value-neutrality question* (cf. [Schurz 1998a](#) , [2024b](#)). Representatives of the empirical-analytical program are (unfortunately) the only ones striving for a clear division between description and value judgments. In hermeneutic approaches, by contrast, descriptive statements and value judgments are often deliberately not distinguished from each other (cf. [Hilgendorf 2000](#) , 39ff; [Dupré 2007](#)), and in critical-dialectical approaches, certain emancipatory norms are presented in a quasi-scientific make-up.

Our analysis has straightforward consequences on some debates on the role of values in science ([Kincaid et al. 2007](#) ; [Carrier et al. 2008](#)). A significant number of contributors in this debate have criticized the ideal of value-free science. Their objections can be resolved by our analysis of means-end inference and the resulting requirement of separating empirically confirmed knowledge from value assumptions. Dupré ([2007](#) , 30–35), for example, argues that science needs to employ value-loaded terms in order to be practically relevant. Dupré is right insofar as the derived norms recommended by scientists employ value-loaded terms. But he seems to overlook that the scientist (qua scientist) recommends these derived norms only *hypothetically* and instrumentally, as means for certain fundamental norms that are given by society.

As emphasized in [prop. 2.5-1](#) , the requirement of value neutrality applies only to the scientific *context of justification* (CJ): it must be independent of fundamental, extra-scientific value assumptions. By contrast, the *context of origination* (CO) inevitably involves extra-scientific values that influence the selection of research questions. However, proponents of scientific value-involvement – such as Douglas ([2000](#)) – argue that even the CJ is influenced by extra-scientific values. In support of this position, the following argument is often made: First, the formulation of the research question in the CO is guided by non-epistemic interests. Second, in the CJ, the research question significantly shapes the choice of scientific method. This shows that non-epistemic interests can have an *indirect* influence on methodological choices, or more broadly, that choices within the CO can have an impact on the CJ.

The argument of indirect influence is certainly correct, but may create confusion. Once a research question is precisely specified, the choice of the optimal method for finding an insightful answer becomes a *purely epistemic* matter. As shown in Schurz ([2024b](#) , 164–167), this also applies to the determination of the “significance level” that is often-cited in this context. In Parker's ([2024](#)) words, the specification of the research question projects an initially non-epistemic interest into a purely epistemic research task. We illustrate this with an *example* : A research project on effective treatments for depression may be guided by non-epistemic interests of the pharmaceutical industry. As a result, the research question may be limited to pharmaceutical interventions: Which antidepressant drugs produce the best outcomes (operationalized in terms of efficiency, sustainability, etc.)? For a research question that has been specified in this precision, the choice of the optimal method (data collection, statistical procedures, etc.) no longer depends on non-epistemic interests. This holds true even if it turns out that, for certain types of depression, psychological therapies are more effective than antidepressant medication. In that case, a properly conducted study would simply yield a largely negative result – that is, antidepressants would be shown to have only limited effectiveness. ⁵

The example of pharmaceutical depression studies leads to another important insight. In our example, it becomes evident in the CJ that restricting the investigation of depression therapies to pharmacological treatments impedes research, because it neglects important causal factors that plausibly lie within the psychological domain. In such a case, the selection of research parameters initially made in the CO must remain open to later revision based on findings in the CJ, even if such revisions run counter to the extra-scientific goals of the research project. Otherwise, one can no longer speak of the CJ being independent from external value commitments.

As a final application of our concept of value neutrality, we turn to the so-called *new demarcation problem* ([Wilholt 2009](#) ; [Holman/Wilholt 2022](#)). This challenge arises for all opponents of value neutrality. The challenge consists in the fact that, on the one hand, advocates of value-laden science massively criticize the capitalist value bias in industry-sponsored studies, while at the same time maintaining that value influences on scientific knowledge are both inevitable and desirable. But this implies that they should regard not only their own (typically left-leaning) values as legitimate, but also those underlying industry-funded

capitalist research (cf. also [Henschen 2021](#) , 20). Most proponents of value-laden science seek to avoid this implication – but how? Some authors have proposed resolving the new demarcation problem by suggesting that scientists should rely on a specific set of preferred values – namely, those reflecting democratic majoritarian values forming the normative foundation of our society. ⁶

Le Bihan ([2024](#)) has raised a fundamental objection to the thesis of democratic value-dependence: modern democratic societies are polarized, with left- and right-leaning segments of the population holding opposing value preferences on many issues – from immigration to climate policy. In this context, the proposal to ground science in “democratic values” offers little guidance, as it leaves open on which side of the political spectrum these “democratic values” are supposed to lie. Moreover, the proposal carries a disturbing implication: in times of political upheaval – such as the current situation in the United States – major scientific findings and textbooks would have to be rewritten, since they would depend on democratic majority values that have now fundamentally shifted. Fortunately, the prevailing conception of science in modern democratic societies has largely freed itself from such political entanglements – and it should remain that way.

From a value-neutral perspective, the question of the “right values” is not a fundamental problem, even in times of value polarization. This is because the goals to which scientists relate their recommendations for means are determined by society. If society is in a state of *value polarization* , then value-neutral policy advice requires conducting hypothetical cost-benefit analyses for *multiple, diverging* value preferences that reflect the political positions of relevant segments of the population (cf. [Carrier 2022](#)). This was also a key element in the learning process reflected in the recommendations of the German Leopoldina commissions during the COVID-19 pandemic (see [Wiesing et al. 2021](#) ; [Schurz 2024b](#) , 158f).

2.10.2 The Argument of Inductive Risk

A much-debated objection against scientific value neutrality is the argument from inductive risk (AIR), originally proposed by Rudner ([1953](#)) and further developed by Douglas ([2000](#)). The AIR assumes that scientific hypotheses are typically supported by empirical evidence only with a certain degree of probability. When scientists accept a hypothesis H as true, they thereby incur a certain “inductive risk” – the risk of error. How small must H’s risk of error be in order for it to be reasonable to accept H as true – that is, to rely on H and be willing to act on the basis of H? The *risk* associated with accepting a hypothesis H depends on two factors: (i) the *error probability* of H (i.e. the probability of $\neg H$ given the total evidence) and (ii) the practical *cost* incurred by accepting H when H is false. While (i) is an epistemic, (ii) is a non-epistemic question, depending on extra-scientific values. H’s total risk of error is given as the product of (i) and (ii). It follows that for hypotheses with practical relevance, their acceptability as a basis for practical action depends inevitably on extra-scientific values.

Quite a few authors see the AIR as convincing evidence for the influence of non-epistemic values on the CJ ([Holman/Wilholt 2022](#) , [Brown 2024](#)). In reality, however, the AIR is weaker than it sounds – a “storm in a teacup.” The AIR can be defused by the demand of *uncertainty transparency* in science, that is, by explicitly stating the error probabilities of scientific claims. This demand not only serves the goal of value neutrality but also helps prevent problematic consequences of expertocracy. We illustrate this with the example of the prediction of *earthquake safety* . Let H be the hypothesis that (in a given region over the next few days) no earthquake of magnitude greater than 6.5 on the Richter scale will occur. Even if the error probability of this hypothesis is only 5%, many affected people would be poorly advised to act on it – at least those whose lives would be seriously endangered in case of error, because their houses are poorly built and at risk of collapse at this level of seismic activity. It is therefore undoubtedly correct that the question of whether to accept non-H rather than H – and thus whether or not to evacuate – depends on non-epistemic cost-benefit considerations (for detailed calculation s. [Schurz 2024b](#) , 160–167 and [sec. 5.9.5](#)).

But does this really mean that scientists should make such value judgments on behalf of the affected people, as proponents of the AIR would have us believe? Certainly not. Instead, the error risks involved should be explicitly stated and explained. The L’Aquila earthquake illustrated this in a tragic way: In October 2012, six earthquake experts and a government official in Italy were convicted in the first instance and sentenced to several years in prison because they had failed to predict an earthquake that occurred on April 6, 2009, in L’Aquila, in which more than 300 people died – instead, they had given the all-clear ([Cocco et al.](#)

[2015](#)). The seismologists defended themselves against the verdict by pointing out that they had not issued a categorical all-clear, but rather a more cautious probabilistic judgment, namely that the observed slight increase in seismic activity was *not significant* , with an estimated error probability of 5%. As a result, the verdict was overturned in 2014; only the government official who had communicated the “all-clear” message to the media was ultimately convicted.

The AIR is based on the mistaken belief that it is the task of science to provide the public with categorical knowledge claims. Quite the contrary: objectively value-neutral science requires that for practical uncertain knowledge claims, the estimated error probabilities should always be reported – an insight first emphasized by Jeffrey ([1956](#) , 237). Referring to Schurz ([2013](#)), both Betz ([2017](#)) and Henschen ([2021](#)) have likewise proposed to solve the problem of inductive risk by making error probabilities explicit. Douglas and like-minded authors attempt to refute this position by arguing that people expect statements from the “authority of science” with clear-cut practical implications – such as giving an “all-clear” ([Douglas 2000](#) , 563; [2009](#) , 135; [Wilholt 2009](#) , 94). But scientists must not give in to such false expectations; rather, it is part of their responsibility to enlighten the public about the limits of the knowable and the risk of error.

However, the critics are right to point out that a statistical significance statement, hardly understandable for laypeople, is not sufficient for such public education. We suggest that error probabilities should be made comprehensible to non-experts by means of *hypothetical cost-benefit considerations* . In the earthquake example, the 5% error probability could be explained as follows: *If* the negative consequences of an earthquake occurring while one remains in their home are estimated to be more than 20 times greater than the costs of a temporary evacuation, *then* it would be advisable to evacuate.

2.10.3 Meta-ethical Justification of Value Neutrality and the Difference between Experience and Valuation

The value-neutrality requirement is based on the thesis that fundamental norm and value sentences are not capable of being justified by empirical science. In these subsections, I discuss this thesis in the context of *meta-ethics* , which is that part of ethics that investigates the meaning and justifiability of ethical statements. First and foremost, value neutrality requires the separation of descriptive statements from value and normative judgments. In formal scientific languages, this separation is already ensured by the differing forms of expressions. However, in everyday language, a problem arises: many terms in our everyday vocabulary carry *both* factual and evaluative content, often closely intertwined, making conceptual separation difficult. Some critics of value neutrality, such as Putnam ([2002](#) , ch. 2) and Dupré ([2007](#)), have claimed that in the case of so-called thick concepts such as “cruelty” or “rape,” such separation is impossible. Yet, as Hare ([1981](#) , 74f) has already argued, this separation is more a *matter of will* than of ability, and with the help of suitable linguistic tools, it is always possible.

The standard meta-ethical justification for the value-neutrality requirement is the *fact – value* or *is – ought* dichotomy , which goes back to David Hume and asserts that value or normative statements cannot be logically or analytically derived from purely descriptive statements ([Hume 1739](#) /1740, 177f). Attempts to prove this prima facie highly plausible thesis by means of modern logic were confronted with unexpected problems, such as the paradox of Prior ([1960](#)) (cf. [Stegmüller 1979a](#) , 183–186). As a result, the dichotomy thesis was called into question by a number of philosophers ([MacIntyre 1981](#) ; [Putnam 2002](#)) and could only be placed on a solid foundation through more recent logical investigations. ⁷ Schurz ([1997](#)) proves that it is only possible to derive nontrivial ethical statements from descriptive premises if one already presupposes so-called *is-ought bridge principles* in one’s ethical theory, such as the utilitarian principle “Whatever is desired by the majority of humans, is ethically good, or ought to be realized.” Is-ought bridge principles are not analytically valid, but are synthetic assumptions belonging to one’s accepted ethical system and are typically controversial. Therefore, norms or values cannot be justified by descriptive (non-ethical) premises without already assuming hidden ethical principles.

There exists a variety of different meta-ethical positions (e.g. utilitarianism and a priorism), but none of them stands in a substantial conflict with Hume’s Is-ought thesis. ⁸ The only exception is the position of so-called *value empiricism* (cf. [Czaniera 2001](#) , ch.1–2). This position assumes a special kind of “value experience,” which, although distinct from biological sensory experience, is analogous to it in that it relates to individual experiences and can serve as a basis for testing in the value sciences. This position has been

advocated by, among others, Firth ([1952](#)), Weingartner ([1978](#) , §7.15 ff), and more recently by Anderson ([2004](#)). In the following, I will use two arguments to show that such “value experiences” behave quite differently from genuine sensory experiences: the normal conditions argument and the independence argument.

A The *normal conditions argument* says that intersubjectively deviating perceptual judgments can almost always be traced back to *defects* in the normal conditions of the observation. But this does almost never apply to intersubjectively deviating value judgments. Let us consider, for example, the presumably true observation sentence: “There is a fire there.” The thesis of the *intersubjectivity* of such an observation sentence implies that any or almost any person would agree with this sentence, provided the *normal conditions* of observation explained in [sec. 2.9.2](#) are fulfilled. If a person within visual range of the fire would really repeatedly and persistently claim “but I see no fire,” then we would have every reason to assume that there is a defect in the person’s normal conditions, and we would be able to find out this defect by independent tests. Such a person would most probably be blind or brain-damaged, or could be in a state of extreme hallucination. For this reason, we would also be entitled to try to *cure* such a person of their observational defect.

But now let us assume that an adult person, under the same normal conditions of observation, systematically *disagrees* with a basic value sentence such as “It is good that *this* person is being kept from dying by being fed,” although we assume that 99% of all persons do agree with it. Then we would *not be entitled* to assume that there is a defect in the ‘normal moral condition’ of this person, that is, that this person is morally inferior and should be cured, just like the blind should be cured of their blindness. Rather, we must *respect* this person’s position *prima facie* as a deviating value position and also – as far as it does not infringe the rights of others – accept it. It is normal in Inuit society, for example, not to give food to a person who is so weak with age that they cannot feed themselves, but to let them die instead, because this is seen as a natural sign that their spirit wants to leave this world. Should we therefore declare the morals of the Inuit to be ‘morally defective,’ and try to ‘cure’ them of their defects? If we treated basic value sentences exactly like observation sentences, we could find ourselves moving dangerously close to medieval religious moral practices.

I do not want to deny with this thesis that there are indeed cases of *psychopathological* blindness of conscience, which are in need of a psychological cure – but these cases are not the *rule*. An extreme claim such as “I love killing other people” would rightly cause anybody to react by thinking “this person has a defect” – but this situation applies only to *very few* deviating basic value sentences, whereby the analogous situation in the case of deviating judgments of perception is the *rule*. Quite in this sense, Prim and Tilman ([1979](#) , 93f) report an experiment on the objectivity of the observation sentence: “there is a piece of chalk in front of me,” which they conducted on their students: the students reacted very indignantly, and declared that dissent in such questions was a problem for an *optician*, and not one of philosophy.

B The *independence argument* goes back to Toulmin ([1950](#) , 127) and runs as follows: Our ‘value experiences’ are *dependent* on our *moral background assumptions*. Our perceptual impressions, on the other hand, remain the same, even when our descriptive background assumptions change radically. For example, we see the Sun make its way across the sky day in, day out. Let us assume that up to now we have held the geocentric view (the Sun revolves around the stationary Earth); but now, after detailed study of Kepler and Newton, we have changed to the heliocentric view (the Earth revolves around the stationary Sun). There has now been a radical change in our background theory – but as already explained in [sec. 2.9.1](#) (contra-3), our visual experience is changed *in no way* by this fact: just as before, we see the Sun drawing an arch across the sky every day. Our visual experience is independent of acquired background assumptions.

This is quite different in the case of value-based moral sentiments. If I see somebody donating a sum of money to the poor, I may spontaneously have the moral sentiment: *that was a good deed*. If I find out, however, that the donor is a politician about to stand for election, who has invested this donation as part of a calculated plan to win votes, but is otherwise more concerned about the well-being of the rich than of the poor, then a quite different moral sentiment will arise within me: I now see the politician as a *hypocrite*.

The normal conditions and the independence arguments strengthen our thesis that an empirical justification of value or norm sentences in analogy to empirical science is not possible. Value and norm-

based judgments are not based on sensory experiences, but on subjective interpretations of experiences that depend on cultural background values and emotional impressions arising in the context of the perceived situation.

2.11 The Justification of Induction: An Unsolvable Problem?

2.11.1 Is Induction Rationally Unjustifiable?

As mentioned in [sec. 2.7.1](#), Hume's problem of induction consists in the question of how inductive reasoning can be rationally justified. David Hume made it clear to the philosophers of his time that, *first*, inductive inferences cannot be justified by means of *logic*, because it is logically possible that the future (or the unobserved) is radically different from the past (or from the observed). However, Hume went much further: he cast doubt on the rational justifiability of induction as a whole (1748, sec. 4). Induction, *second*, also cannot be justified by observation, because in inductive inferences one infers that which has *not yet* been observed. And *third*, inductive inference also cannot be justified by induction, as this would lead us straight into a *circular argument*: if we think inductive inference is reasonable because it has been corroborated so far, then we are assuming that what has been corroborated so far will – by inductive inference – also be corroborated in the future. In such a circular argument, we assume exactly the thing we want to show, and so we show nothing at all. All three traditional methods of scientific justification – logic, experience, and induction – thus appear to fail when it comes to justifying induction itself. Hume was also the first to show that, *fourth*, a reformulation of the induction problem using probability theory does not help (1748, sect. 6). If we argue that inductive inferences do not always, but at least *with great probability* lead us from true premises to true conclusions, then we are assuming a probabilistic induction principle, according to which the *event frequencies* observed so far may be inductively transferred to the future or to non-observed cases.

These were the reasons which led Hume to his deeply *skeptical* thesis, that inductive inference is not capable of being rationally justified, but is based merely on psychological *habit* (1748, sect. 5). Finding a satisfactory solution to Hume's induction problem would be of the greatest importance. And yet, to this day, there is no generally accepted solution to the problem. ⁹We briefly discuss five solution attempts that have been suggested in the last few decades but could not find general acceptance:

1 *Induction is not needed in science?* Popper's attempt to circumvent the induction program was to argue that scientific methodology doesn't need induction. In [sec. 2.7.2](#), we have seen, however, that Popper's argument fails, because Popper's own methodology of "corroboration" contains in its heart the principle of epistemic induction.

2 *Induction as rational by definition?* Proponents of the analytic account (e.g. [Strawson 1952](#), 257; [Ayer 1956](#), 74f, [Pollock 1974](#), 204) have argued that induction is rational by the very definition or meaning of rationality. Recall, however, that by a rational justification we don't mean a mere 'intuition of rationality' (which is then declared as being 'true by definition'), but a demonstration that inductive inferences are truth-conducive, in the sense of optimizing our success in delivering true predictions. This cannot be shown by a mere 'definition of rationality,' no more than diseases can be cured by a 'definition of health.' Nor is Strawson's claim (ibid., 249) correct that we cannot imagine a situation in which inductive inferences fail (cf. [Salmon 1974](#)). We can not only conceive of such situations in thought (consider *Alice in Wonderland*); in fact, millions of people *do* believe in the superiority and even in the 'higher rationality' of non-inductive 'methods,' be it clairvoyance, fortune-telling, or God-guided inner intuition. So, a satisfactory justification of induction would not only be of high philosophical, but also of high *cultural* importance, as part of the enterprise of explaining scientific rationality.

3 *Rule-circular 'justification' of induction by induction?* The circularity involved in the 'inductive justification of induction' has been called *rule-circularity* ([Ladyman/Ross 2007](#), 75), because here, the conclusion (the reliability of induction) is not an element of the premise set, but is presupposed by the underlying inference rule (namely induction). Several philosophers expressed the hope that rule-circular arguments have epistemic value and might offer a solution to the problem of induction. ¹⁰The following counterexample by Salmon ([1957](#), 46) destroys this hope. It shows that the same type of rule-circular

argument that ‘justifies’ the reliability of induction can be used to ‘justify’ the rule of counterinduction, which says, roughly speaking, that one should predict the opposite of what one has observed:

(2.11-1)

Rule-circular justification of induction :

Premise : Past inductions have been successful.

Therefore, by rule of induction:

Inductions will be successful in the future

Rule-circular justification of counterinduction :

Premise : Past counterinductions have not been successful.

Therefore, by rule of counterinduction:

Counterinductions will be successful the future.

The two circular arguments share exactly the same structure, the premises of both arguments are true, and yet they lead to opposite conclusions. This proves that such rule-circular arguments are epistemically worthless pseudo-justifications. A generalization of this refutation for inferences to the best explanation is found in Douven ([2011](#)), and for arbitrary rule-circular arguments in Schurz ([2024a](#) , [sec. 4.3](#)). Rule-circular justifications can also be given for dangerous fundamentalistic rules, such as the “rule of blind trust in my leader,” abbreviated (BTL), whose rule-circular justification goes as follows: “I believe in the reliability of the rule BTL because this was assured to me by my leader.”

4 Induction justified by uniformity of nature? John Stuart Mill ([1865](#) , III .3.1) and Bertrand Russell ([1912](#)) held the view that the reliability of induction could be justified by assuming the uniformity of nature (i.e. nature’s future resembles its past). The main problem with this approach is that it also leads to a *justificatory circle* , since the assumption of uniformity can itself only be justified by an inductive inference. Note that as in ([2.11-1](#)) above one may give a corresponding circular justification of counterinduction by counteruniformity: counterinduction is justified by counteruniformity which is in turn justified by counterinduction from past observations. A “localized” version of the uniformity-account has been proposed by Norton ([2003](#)). For Norton, inductive reasoning is not governed by formal but by domain-specific or “local” inference rules such as “if some samples of Bismuth melt at 271 ° C, then all samples of Bismuth will do so.” Local inductions are in turn justified by local uniformity assumptions, for example, “samples of the same element agree in their physical properties.” However, Norton’s proposal is subject to exactly the same charge of circularity as the Mill-Russell approach, since local uniformity assumptions must be justified by inductive inferences (cf. [Kelly 2010](#) ; [Schurz/Thorn 2020](#)).

5 Induction justified by inference to the best explanation? Several authors ([Harman 1965](#) ; [Armstrong 1983](#) , sec. 6; [Lipton 1991](#) , 69) have argued that inductive inferences are justified because they are instances of inferences to the best explanation (IBE). Recall from [sec. 2.7.3](#) that an IBE infers from an observed phenomenon that the premises of its best (available) explanation are probably true. An inductive generalization – so the argument goes – is an instance of an IBE , because the best explanation of a regularity observed so far (e.g. “so far all observed ravens were black”) is the corresponding lawlike generalization (“It is a law that all ravens are black”).

The major problem of the IBE-account is, however, that the justification of IBEs is even more difficult than the justification of induction. This is so because IBEs involve even more uncertainties than inductive inferences (recall [sec. 2.7.3](#)). Analogous justification strategies to those discussed above for induction have been proposed for IBEs (cf. [Bird 1998](#) , 171). Armstrong ([1983](#) , 53), for example, has argued that IBE is rational ‘by definition.’ Lipton ([1991](#) , 167ff) and Psillos ([1999](#) , 82) suggested to justify IBE in a rule-circular way. But we have seen that both strategies are bound to failure. Moreover, it is unclear why one should regard the explanation of “All so-far observed Fs have been Gs” by “Because it’s a law that all Fs are Gs” as a *best* explanation, if one does not accept the inductive premise that the world is uniform. If one assumes instead that from time to time the laws of nature undergo radical changes, the explanation “Because it was *so-far* a law that Fs are Gs” seems to be the better explanation.

2.11.2 The Optimality of Meta-induction

Due to the many failed attempts to solve the problem of induction, quite a few contemporary philosophers consider it unsolvable. ¹¹ We do not share this view. We agree with Hume's skeptical diagnosis that no noncircular argument can establish that induction is *reliable*, that is, will be predictively highly successful, whatever the future looks like. However, there is an alternative justification strategy that is not blocked by Hume's argument: namely, the attempt to show that induction is an epistemically *optimal* prediction method, in the sense of being most successful among all prediction methods that are *accessible* to a given agent. Even in radically skeptical scenarios in which all methods of prediction are unreliable, induction can still be optimal in the sense of 'being the best of a bad lot.'

A noncircular optimality justification of induction was developed by Schurz (see, e.g., [Schurz 2008b](#), [2019a](#)). A precursor to this approach is Hans Reichenbach's justification of induction as the "best alternative" ([Reichenbach 1949](#), §91). However, Reichenbach failed to develop a robust optimality argument in favor of induction, as he based his argument on *object-induction*, that is, induction aimed at predicting events. Yet, nothing can rule out the logical possibility of a clairvoyant who predicts a random sequence with a significantly higher success rate than the scientific inductivist (see [Skyrms 1999](#), sec. III.4). By contrast, the approach developed by Schurz is based on the idea of *meta-induction*.

The meta-inductivist applies the inductive method to all accessible prediction methods and attempts to derive an optimal prediction strategy based on their observed track records. The simplest meta-inductive method is *Imitate-the-best*, abbreviated as ITB. It predicts what the actually best non-MI method predicts (resolving ties according to a given ordering of accessible methods). ITB's success will be optimal in the long run if, after some finite time, there is a *unique* best method that ITB imitates forever (cf. [Schurz 2019a](#), [sec. 6.1](#)). Interestingly, ITB is *not* universally optimal. ITB's optimality breaks down when it plays against 'deceiving' methods that lower their success rate as soon as their predictions are imitated ([Schurz, 2008b](#), sec. 4). A realistic example is predictions of stock values in a 'bubble economy:' here, the prediction that a given stock will yield a high return leads many investors to buy that stock, causing it to crash by doing so (since the stock lacks sufficient economic support).

A more refined meta-inductive strategy is attractivity-weighted meta-induction. This meta-inductive method, abbreviated as AW, predicts a weighted average of the predictions of all methods accessible to it, with weights determined by the methods' past success rates (overview in [Schurz 2019a](#), sec. 6–7). Building on results from mathematical learning theory ([Cesa-Bianchi and Lugosi 2006](#)), Schurz ([2008b](#)) could show that AW is indeed universally optimal: in all prediction games with arbitrary sequences of events and a finite number of available methods, AW's predictive long-run success is greater than or equal to the maximal long-run success of all accessible methods ([Schurz 2008b](#), theorem 4; [2019a](#), theorems 6.8–6.11). Even in 'radically nonuniform' scenarios in which the success rates of the leading players permanently oscillate around each other, AW's success rate converges to the maximal success rate (or to a greater value), with a small short-run loss that converges to zero with an increasing number of predictions.

The precise formulation of the meta-inductive optimality theorem runs as follows. Let $(e) =_{\text{def}} (e_1, e_2, \dots)$ be a sequence of events e_n ordered along a discrete time $n = 1, 2, \dots$, whose possible values are coded by real numbers in the interval $[0,1]$. Let $M = \{M_1, \dots, M_m, AW\}$ be a finite set of (m) simultaneously accessible prediction methods together with AW. The methods $M_1, \dots, M_m$ may be given as computational algorithms, real-life experts or even (purported or real) clairvoyants; their "simultaneous accessibility" means that the epistemic agent can track and compare their success records, either by observation or by simulation. The pair $((e), M)$ is called a *prediction game*. In each round n of the game, it is the task of each method to deliver a prediction $\text{pred}_{n+1} \in [0,1]$ of the next event e_{n+1} of the sequence. The deviation of the prediction pred_n from the event e_n is measured by a normalized loss function, $\text{loss}(\text{pred}_n, e_n) \in [0,1]$. The optimality theorem below holds for all *convex* loss functions (a generalization is possible that even holds for all loss functions). The predictive success of a method M at a given time point n is evaluated by its *success rate* $\text{suc}_n(M)$, defined as the sum of M 's scores achieved until time n divided by n (where M 's score achieved at time n is $1 - \text{loss}(\text{pred}_n, e_n)$). The attractivity of a method M for AW at time n , $\text{at}_n(M)$, is defined as $\text{suc}_n(M) - \text{suc}_n(AW)$. The prediction of the meta-inductive method AW for the next time $n+1$, $\text{pred}_{n+1}(AW)$, is the normalized weighted average of the predictions of the methods $M_1, \dots, M_m$ for this time with weights

defined as $w_n(M_i) =_{\text{def}} e^{\eta \cdot n \cdot \text{at}_n(M_i)}$, where $\eta = \sqrt{8 \cdot \ln(m)/(n+1)}$. “ maxsuc_n ” denotes the maximal success rate of the accessible methods at time n . With these definitions, the optimality theorem runs as follows:

(Prop. 2.11-1) *Universal optimality of AW*: ¹²

For every prediction game $((e), \{M_1, \dots, M_m, AW\})$:

- i. (Short run) $(\forall n \geq 1:) \text{suc}_n(AW) \geq \text{maxsuc}_n - 1.77 \cdot \sqrt{\ln(m)/n}$.
- ii. (Long run) $\lim_{n \rightarrow \infty} (\text{suc}_n(AW) - \text{maxsuc}_n) \geq 0$.

In the short run, weighted meta-induction may suffer from a possible loss (or “regret”) compared to the leading method, caused by the fact that AW must base her prediction for the next time on the *past* success rates of the accessible methods. However, proposition 2.11-1(i) states a worst-case upper bound for this loss, which is small if the number of rounds n is not smaller than the number of candidate methods m , and which quickly converges to zero when n grows large.

AW cannot get deceived by a method because it never favors only one method but a weighted average of all attractive methods. A second important feature of AW is that it tends to *ignore* methods that are constantly worse than itself, since exponential weights with negative attractivities quickly converge to zero. This ignorance-tendency is necessary for AW’s long-run optimality and a crucial distinction of AW compared to success-weighted meta-induction that uses simple success rates as weights and is provably nonoptimal (Schurz 2019a, sec. 6.8.1). It follows that in games in which one method dominates its competitors, the weight of the best method will quickly converge to 1, and AW’s behavior will converge to that of ITB.

The optimality result holds independently of any assumptions about the possible worlds or their prior probabilities. This means that meta-induction possesses a mathematical a priori justification. But what does this result imply for the justification of object-induction? In possible worlds that contain genuine clairvoyants, object-induction cannot be optimal; in such worlds, meta-induction does not favor the object-inductivists, but rather the clairvoyants. However, the a priori justification of meta-induction provides us with the following noncircular *a posteriori justification of object-induction*: we know from experience that object-inductive prediction methods have so far outperformed all other available prediction methods. Therefore, we are justified – by meta-induction – in continuing to use object-induction in the future. This justification is, unlike the “inductive justification of induction,” no longer circular, because we have an independent a priori justification of meta-induction.

In the meantime, a lively debate has emerged regarding the prospects of the meta-inductive optimality justification, in which further challenges have been raised and solutions proposed. ¹³In Schurz (2024a), optimality justifications are extended to the entire field of epistemology.

2.12 Introductory Literature and Exercises

2.12.1 Introductory Literature

On the method of rational reconstruction, see Stegmüller (1969 ff, vol. IV.1). A classic on the German value judgment debate is Albert and Topitsch (1971, ed.), for more recent debates, see Kincaid et al. (2007) and Carrier et al. (2008); a recent introduction is Elliott (2022). A collection on the question of the unity of sciences is Rahman et al. (eds., 2004). Psillos (1999) offers an excellent introduction to the empiricism versus realism debate on philosophy of science. Recent overviews on the metaphysics of science are Schrenk (2017) and Hüttemann (2021). Rock (1984) introduces into the cognitive psychology of perception. Swinburne (1974) and Henderson (2018a) offer introductions to the problems of induction. Lipton (1991) is a classic on inference to the best explanation and Niiniluoto (2018) is a good introduction to abduction in science.

2.12.2 Questions and Exercises

Sec. 2.1

1. What is meant with “context of discovery” versus “context of justification?”
2. Explain the normativist argumentation and its criticism.
3. What is the supreme epistemic goal of science? Does it presuppose metaphysical realism?
4. Explain the method of rational reconstruction.

Sec. 2.2–2.4

1. Explain the five common epistemological assumptions of the sciences.
2. In which sense are minimal realism and minimal empiricism “minimal?”
3. What is the difference between falsifiability and fallibility?
4. Explain the common methodological features of sciences.
5. To what extent are the concepts of “natural sciences,” “human and social sciences,” and “humanities” ambiguous?
6. How do empirical, experimental, and dissecting sciences differ from one another in regard to the degree of their empirical invasiveness? Classify educational science, human medicine, and cultural studies within this classification.

Sec. 2.5–2.6

1. Explain the means-end inference using examples from your area of study.
2. Explain the suggested formulation of the value-neutrality thesis and all concepts contained in it.
3. Discuss: How is the requirement of “value neutrality” related to that of “political correctness?” Are the two in conflict?
4. Discuss physical reductionism and its criticism. What is “supervenience?”

Sec. 2.7

1. Which kind of inference (deductive, inductive, or abductive) is being used in the following examples?
(i) Most light switches are working. I am pressing the switch. Therefore, the light will go on. (ii) All humans are mortal. Aristotle is human. Therefore, Aristotle too is mortal. (iii) My refrigerator has always worked well up to now. Therefore, I will be able to rely on it tomorrow. (iv) Every time Gregor talks about his brother, his facial expression becomes strained. Therefore, he has a problem with his brother.
2. Explain the principles of methodological, logical, and epistemic induction. What are the Popperian arguments against the first two kinds of induction?
3. Give an example of an inference to the best explanation.
4. Which inferences in 1 are monotonic and which are non-monotonic?

Sec. 2.8–2.9

1. What is the difference between metaphysical and hypothetical-constructive realism?
2. Explain the argument of radical constructivism against realism and its criticism.
3. One person could fail to notice another because (a) he is deliberately ignoring her, (b) he is blind, (c) he is asleep, or (d) it is pitch dark. Which normal condition of observation is infringed in each case?

Sec. 2.10–2.11

1. Explain Habermas’ threefold division of scientific disciplines according to their respective interests. How is this view criticized?
2. Discuss the normal conditions and the independence argument.
3. Recapitulate David Hume’s four arguments against the possibility of justifying induction.

4. Explain the structural equality of the inductive justification of induction and the counterinductive justification of counterinduction.
5. What is a meta-inductive prediction method? In which sense is attractivity-weighted meta-induction optimal?

Notes

1. Cf. Oppenheim/Putnam ([1958](#)), Causey ([1977](#)), Hofer/Smeenck ([2016](#)); Hüttemann ([2021](#) , ch. 5–6). [↩](#)
2. Cf. Humphreys ([1997](#) , 10), Ladyman/Ross ([2007](#) , sec. 1.6), and Hüttemann ([2021](#) , 151–155). [↩](#)
3. For arguments in this direction cf. Chalmers ([1986](#) , [sec. 3.4](#)), van Fraassen ([1980](#) , 56ff), Albert ([1985](#) , sec. II.8), Brewer/Lambert ([2001](#) , S182f). [↩](#)
4. For example, Putnam ([1962](#)), Shapere ([1982](#)), Kuipers ([2000](#)), Giere ([1999](#)), Adam ([2004](#)). [↩](#)
5. It is well known that the so-called *positivity bias* in scientific research often leads to the non-publication of negative results, which causes significant distortions in the research record ([Ioannidis 2005](#) ; [Schooler 2011](#)). One example is evidence of bias in pharmaceutical studies in favor of their industrial sponsors ([Brown 2008](#)). However, such distortions are not an inherent feature of science but rather a *defect* of scientific practice under external pressures. In response, a number of scientific journals have introduced measures to ensure the publication of negative research findings. [↩](#)
6. For example, Elliott ([2017](#)), Schroeder ([2021](#)), Alexandrova/Fabian ([2022](#)). [↩](#)
7. S. Pigden ([1989](#)), Schurz ([1997a](#)), Russell ([2022](#)); overviews in Hudson ([1969](#)) and Pigden ([2010](#)). [↩](#)
8. Overview of meta-ethical positions is, for example, found in Frankena ([1963](#)), Kutschera ([1982](#) , sec. 2), Schurz ([1997a](#) , sec. 11.3), and Sayre-McCord ([2023](#)). [↩](#)
9. Cf. Stegmüller ([1977](#) , vol.II, ch. 4), Swinburne ([1974](#)), Rescher ([1980](#)), Howson ([2000](#)), Ladyman ([2002](#) , ch. 1), Schurz ([2019a](#) , ch. 2–4). [↩](#)
10. For example, Braithwaite ([1974](#)), Black ([1974](#)), van Cleve ([1984](#)), Papineau ([1993](#) , sec. 5), Goldman ([1999](#) , 85), and Psillos ([1999](#) , 82). [↩](#)
11. See, e.g. Stegmüller ([1977](#) , vol. II, ch. 4), Greco ([2005](#) , 265); Williamson ([2023](#) , 451). [↩](#)
12. The loss term in (i) simplifies the loss term in th. 6.9(ii) of Schurz ([2019a](#)), since $\sqrt{2} + 1/\sqrt{8} < 1.77$. [↩](#)
13. See, for example, Arnold ([2010](#)), Henderson ([2018a](#) , sec. 7.3), Sterkenburg ([2019](#)), Henderson ([2023](#)), Douven ([2023](#)), Briesen ([2026](#)), Shogenji ([2023](#)), Wolpert ([2023](#)), Pitts ([2023](#)), Kimling ([2023](#)); Replies in Schurz ([2021b](#) , [2023a](#) , 2023b, 2024c, 2025). [↩](#)

3 The Conceptual Toolkit

Language, Logic, and Probability

DOI: [10.4324/9781003675075-3](https://doi.org/10.4324/9781003675075-3)

This chapter has been made available under a CC BY-SA license.

3.1 Kinds and Classification of Concepts

The basic logical building blocks of any language are its concepts. A concept is understood here as a linguistic expression *together* with its meaning. Grammatically, concepts are *beneath* the level of the sentence – they are *subsential*. “Grass” and “green” are concepts; “grass is green” is a sentence. Though concepts signify something, they cannot in themselves be true or false; only sentences composed of concepts can. Concepts may be *atomic* or *complex*; for example, “grass” is an atomic, while “green grass” is complex. Sentences are composed of concepts, and the nature of a sentence depends on the nature of its concepts. So, we will *first* deal with the various different kinds of scientific concepts, before dealing with the various different kinds of scientific sentences. Concepts may be classified in three respects: first according to their logical type, second according to their content type, and third according to their gradation (or scale) type.

3.1.1 The Logical Type of Concepts

Logical grammar classifies concepts according to their logical-semantic function. [Fig. 3.1-1](#) presents an overview.

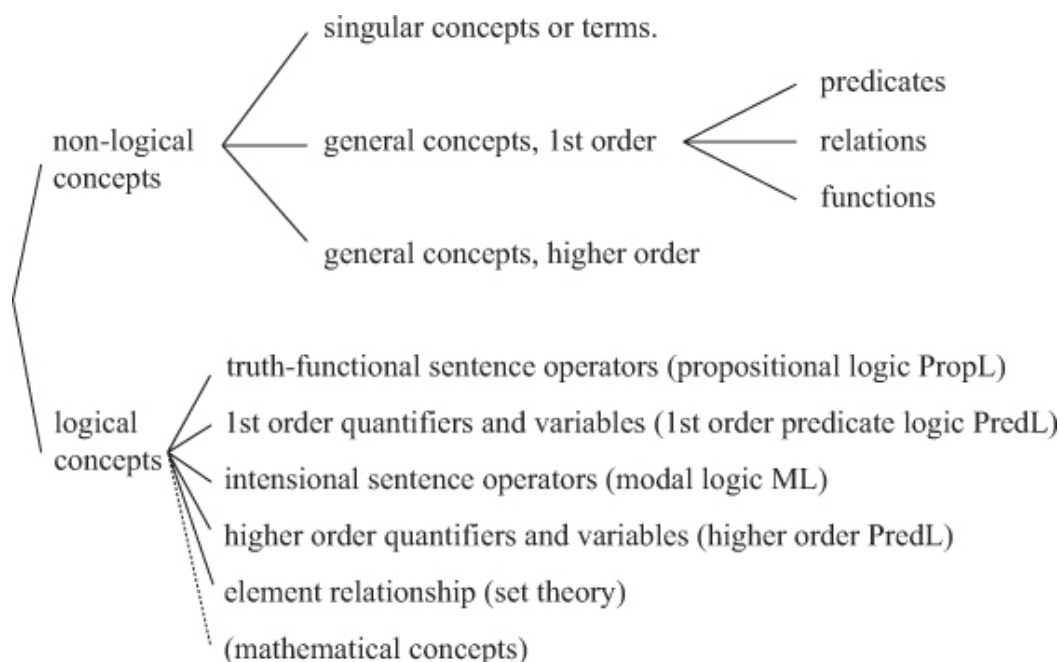


Figure 3.1-1 Division of concepts according to their logical type. [↩](#)

1. *Nonlogical concepts* denote or signify something in the world . By contrast, logical concepts have a purely *structural* function.

1.1. *Singular concepts* or *terms* always designate a single *individual* (or spatiotemporal location). Singular terms subsume (a) *proper names* (e.g. “Saul Kripke,” “Salzburg”), (b) *ostensive* or *indexical terms* (e.g. “I,” “now,” “here,” and “that person there”), and (c) *definite descriptions* (“that x which has characteristic property F”). *Formal notation*: atomic singular terms are called *individual constants* and are formalized by small letters $a, a_1, a_2, \dots, b, \dots$

1.2. *First-order general concepts* are applied to singular terms: this application is called *predication* , and through it, an *atomic sentence* is generated.

1.2.1. *First-order predicates* denote features which can be true of individuals or not. One may distinguish between *property concepts* such as “x is big,” and *kind concepts* such as “x is a person.” *Formal notation*: predicates are formalized using capital letters $F, G, \dots$ (or $F_1, F_2, \dots$). Predicates are *one-place* , that is, they have one argument *place* , which is written as a *variable* after the predicate letter: “Fx” for “x is an F.” For example, if “Fx” stands for “x is slim,” and “a” for “Peter,” then “Fa” stands for the sentence “Peter is slim.”

1.2.2. A (first-order) *n* -place *relation sign* denotes a relation between *n* given individuals. *Formal notation* by uppercase letters $R, Q, \dots$, for example, if “Rxy” stands for the two-placed relation “x is a brother of y,” “a” for “Peter,” and “b” for “Paul,” then “Rab” stands for the sentence “Peter is a brother of Paul.”

1.2.3. *Function signs*: a one-place (or *n*-place) function sign denotes a function which uniquely assigns to one (or *n*) individuals a further individual. *Formal notation* usually by lowercase letters $f, g, \dots$. For example, if “f(x)” stands for “the mother of x” and “a” for “Peter,” then “f(a)” is a singular term designating Peter’s mother. In mathematics, functions play a fundamental role as operations on numbers; for example, addition + is a two-place function, which assigns to the respective numbers *x* and *y* their sum $+(x,y)$, in infix notation written as “ $x+y$.” Moreover, all *quantitative* (e.g. physical) *properties* of individuals are functions: for example, $m(a)$ denotes the *mass* of the body *a*, and “ $m(a) = 5.6 \text{ kg}$ ” stands for “the mass of *a* is 5.6 kilograms.”

1.3. *Second (or higher)-order predicates and functions* are applied to sentences or predicates and thereby generate sentences. An example of a higher-order predicate is the concept of *truth* which applies to sentences.

2. *Logical concepts*: there are different kinds of logical concepts, and their principles are explicated in corresponding kinds of logics.

2.1. *Truth-functional sentence operators – Propositional logic PropL* (where $A, B, \dots$ stand for arbitrary propositions):

- the negation $\neg$ (read: $\neg A$ – *not A*),
- the conjunction $\wedge$ (read: $A \wedge B$ – *A and B*),
- the disjunction (the *inclusive* or) $\vee$ (read: $A \vee B$ – *A or B or both*),
- the so-called ‘material’ implication $\rightarrow$ (read: $A \rightarrow B$ – *if A, then B*),
- the ‘material’ equivalence $\leftrightarrow$ (read: $A \leftrightarrow B$ – *A if and only if (iff) B*),
- *convention*: “ $\leftrightarrow_{\text{def}}$ ” stands for “equivalent by definition”; similarly for “ $=_{\text{def}}$.”

Further sentence operators are thereby definable. For example, the exclusive *either-or* $\dot{\vee}$ is defined as $A \dot{\vee} B \leftrightarrow_{\text{def}} (A \vee B) \wedge \neg(A \wedge B)$ (*A or B, but not both*). The logical laws of these sentence operators are explicated by (non-modal) propositional logic.

A sentence operator, when applied to given sentences, generates a new, complex sentence. A sentence operator is called *truth functional* (or *extensional*) iff the truth *value* of the complex sentence is uniquely determined by the truth values of its component sentences (arguments). The truth-functionality of the sentence operators of PropL is expressed by their “truth tables”: every sentence *A* is either true or false; $\neg A$ is true iff *A* is false (else false); $A \wedge B$ is true iff both *A* and *B* are true (else false), etc. *Note*: the material implication $A \rightarrow B$ is defined as true iff *A* is false or *B* is true, that is, $A \rightarrow B \leftrightarrow \neg A \vee B$ is valid. The material implication is *weaker* than the natural language if-then, because its truth does not presuppose any connection between the contents of *A* and *B*.

2.2 *First-order quantifiers and variables – predicate logic PredL* : first-order *quantifiers* quantify over individuals with help of first-order *individual variables* $x, y, \dots (x_1, x_2, \dots)$. The two most important kinds of quantifiers are:

- the *universal quantifier* $\forall x$: read “ $\forall xFx$ ” as “For all x : Fx ,” and
- the *existential quantifier* $\exists x$: read “ $\exists xFx$ ” as “For at least one x : Fx .”

The class of individuals to which a quantifier refers is called its (semantic) domain D . So $\forall xFx$ says that all individuals in D have property F , and $\exists xFx$ that some individual in D has property F . When nothing further is stated, D is identified with the *universal* domain, that is, the class of all individuals in the given scientific area.

We say that the individual variable “ x ” is *bound* by the quantifier “ $\forall x$ ” or “ $\exists x$.” The subformula “ Fx ” in “ $\forall xFx$ ” is called the scope of the quantifier “ $\forall x$.” Formulas in which all individual variables are *bound* are called *closed formulas* or *sentences*. The expression “ Fx ” (for “ x is an F ”) is called an *open* formula, because the individual variable “ x ” is *free*, that is, not bound by any quantifier. Open formulas such as “ $Fx \wedge Gx$ ” express complex concepts.

A distinguished logical relation is *identity* “ $x = y$,” standing for “ x is identical with y ,” that is, “ x ” and “ y ” denote the same individual. The logic of sentences which are made up of singular terms, general first-order concepts, extensional sentence operators, and first-order quantifiers, is called first-order predicate logic PredL (introduction see [Klenk 1989](#) , [Copi et al. 2010](#)). First-order PredL contains PropL as a fragment and is one of the most important formal frameworks of philosophy of science, among other reasons because type-free set theory, upon which all of classical mathematics builds, is expressible in this framework.

2.3. *Intensional sentence operators* are not truth functional. An example is the necessity operator $\Box$ (“ $\Box A$ ” for “ A is necessary”): the truth value of “ $\Box A$ ” is not determined by the truth value of A . For instance, if A and B stand for the sentences “I have brown hair” and “ $2 + 2 = 4$,” respectively, then both A and B are true, but only $\Box B$ is true, while $\Box A$ is false, as the fact that I have brown hair is not necessary. Other examples are *deontic* sentence operators (“ A is obligatory”). The logic of intensional sentence operators is *generalized modal logic* (see [Hughes/Cresswell 1996](#)).

2.4. *Second (or higher) order quantifiers* quantify over states of affairs or properties and are dealt within second (or higher) order PredL. The logic of higher-order quantifiers is called logical *type theory*, which goes back to Russell ([1903](#)).

2.5. *Set-theoretical symbols*: sets are aggregations of individuals into a complex individual called a “set.” – *Formal notation*:

- “ $x \in y$ ” stands for “ x is an element of the set y .”
- $\{x: Fx\}$ stands for the set of all individuals x which possess the property F .
- $\{a_1, \dots, a_n\}$ stands for the set consisting of the individuals $a_1, \dots, a_n$.
- $(a_1, \dots, a_n)$ by contrast, stands for the ordered sequence consisting of $a_1, \dots, a_n$.

Note: sets are *invariant* with respect to (w.r.t.) permutations or repetitions of their elements, hence $\{a, b\} = \{a, a, b\} = \{b, a\} = \{b, a, a, b\}$ (etc.). By contrast, sequences or n -tuples consist in an ordering of individuals that allows for repetition: that is, $(a, b, b, c) \neq (a, b, c) \neq (b, a, c)$.

Set theory is a part of *informal* logic or metalogic, as “ $\in$ ” too is a logical concept in a broader sense. The axioms of type-free set theory (after *Zermelo* and *Fraenkel*) are expressed in the language of first-order PredL, but they go *beyond* purely logical principles (for an introduction, see van [Dalen et al. 1978](#) , [Hrbacek/Jech 1999](#)).

- “ $x \subset y$ (resp. $x \subseteq y$)” stands for “ x is a proper (resp. a proper or improper) subset of y .”
- “ $f: A \rightarrow B$ ” stands for “ f is a function which assigns to every individual x in A exactly one individual y in B ”.

Further set-theoretical notions are explained in the text.

2.6. *Mathematical concepts* , as with logical concepts, characterize *abstract-conceptual* objects or functions. They may be characterized either by internal axioms of mathematical theories or by set-theoretical definitions. Among the most important mathematical concepts are concepts for different kinds of numbers (e.g. $\mathbb{N}$ for the set of natural and $\mathbb{R}$ for the set of real numbers), and function concepts expressing operations on numbers (addition, multiplication, etc.).

3.1.2 Syntax and Semantics, Statements, and Models

Since Morris ([1946](#)), the dimensions of language are subdivided into *syntax* , *semantics* , and *pragmatics* . Syntax deals with the grammatical form of linguistic expressions, semantics with the meanings of expressions, and pragmatics finally, with those properties of expressions which are determined by their *context* . In this section, we outline some foundations of syntax and semantics.

From atomic concepts, one may construct terms and formulas of arbitrarily high complexity. The translation of a natural language sentence into a logical sentence is called *formalization*. Here are some examples:

(3.1-1) *Examples of formalizations:*

Translations: a – this animal, Rx – x is a raven, Bx – x is black.

- i. This animal is a raven and it is black. Formalization: $Ra \wedge Ba$.
- ii. All ravens are black. *Partial formalization:* For all x: if x is a raven, then x is black. *Formalization:* $\forall x(Rx \rightarrow Bx)$.
- iii. Some ravens are not black. *Partial formalization:* There exists an x which is a raven and is not black. *Formalization:* $\exists x(Rx \wedge \neg Bx)$.

Sentences with the same natural language syntax can have a very different logical syntax. In particular, the natural language copula “is” serves three different logical functions:

(3.1-2)

<i>Sentence</i>	<i>Formalization</i>	
	Ta	“is” as predication
i. Peter (a) is tall (T).	$a = f(b)$	“is” as identity
ii. Peter is the father (f) of Paul (b).	$\forall x(Rx \rightarrow Bx)$	“is” as inclusion of properties
iii. A raven (R) is a bird (B).		

In ([3.1-2](#))(iii), there is a *concealed* universal quantifier, which is revealed by formalization. The formalization of ([3.1-3](#)) reveals a *concealed* existential quantifier:

([3.1-3](#)) Peter (a) has a son. *Partial formalization:* There exists an x, such that x is the son (S) of Peter. *Formalization:* $\exists x Sxa$.

Logical formalization is an important aid to the advanced philosophy of science, to uncover the logical structure of sentences, and building on this, to discover logical connections. In this book, formalization is used rarely and as little as possible; for most purposes, partial formalizations such as in ([3.1-1](#)) above are sufficient.

In *semantics* , one differentiates between two kinds of semantic denotation: every expression has on the one hand a *meaning* , and on the other an object *reference* . (Terminologies differ, however; cf. [Kutschera 1976](#) , ch. 2; [Morris 2006](#) .) The meaning or intension of a concept is that which must be known to a speaker in order to *understand* the concept properly. The meaning of an expression *together* with the *facts* of the world determines the expression’s reference, that is, the real entity denoted by the expression (if there is no such entity, the reference is empty). There are controversial views on the question of wherein *exactly*

meaning exists (in the imagination, in thought, in the definition, or in usage?). The reference of a singular term is uncontroversially an individual. The reference of predicates is also controversial: some authors (e.g. [Carnap 1956b](#), ch. I) identify the reference of a predicate with its *extension*, that is, the *set* of individuals which fall under the predicate; other authors (e.g. [Fodor 1990](#), 93) identify the reference of a predicate with the *real property* or natural kind denoted by it.

The meaning of sentences is usually identified with the *proposition* expressed by the sentence. According to the traditional view, logically or analytically equivalent sentences express the *same* proposition. The reference of sentences is sometimes identified with truth values (e.g. in [Frege 1892](#)), and sometimes with events or real states of affairs. The following definition of “state of affairs” and “fact” is uncontroversial: every descriptive sentence expresses a *state of affairs*; and if the sentence is true, the state of affairs really exists and is then a *fact*.

In *logical semantics*, which goes back to *Alfred Tarski* (1936a), one assigns to linguistic expressions formal object references, in the form of set-theoretical *interpretations*. Informal set theory functions in this case as a *metalanguage*, in which the semantic relationship between the *object language* and a hypothetical object structure is modeled. The concrete definition is as follows: an extensional *interpretation* of a logical language $\mathcal{L}$ is a pair (D, I) , consisting of a domain of individuals D and an interpretation function I , which assigns to every individual constant a and variable x from $\mathcal{L}$ an individual $I(a)$ in D , to every n -placed relation sign R in $\mathcal{L}$ a subset $I(R)$ from D^n (where D^n is the set of all n -tuples of D -individuals), and to every n -placed function sign f in $\mathcal{L}$ a function $I(f): D^n \rightarrow D$. An extensional interpretation (D, I) makes an atomic formula Fa true, in short $(D, I) \models Fa$, iff $I(a) \in I(F)$. This basic truth definition is recursively extended to arbitrary formulas as usual (cf. [Machover 1996](#), sec. 8.4; [Bergmann et al. 1998](#), sec. 8.7).

The extensional interpretation of an object language can also be seen as a *relationship of translation* between a logical and a set-theoretical language – for example as follows:

(Prop. 3.1-1) *Translation (I) between logical and set-theoretical language*, where $I(a) = a$, $I(F) = F$, etc.:

Logical sentence	Set-theoretical sentence (translation)
Fa	$a \in F$
$\forall x(Fx \rightarrow Gx)$	$\forall x(x \in F \rightarrow x \in G)$, in short: $F \subseteq G$
$\psi(F)$	$F \in \psi$ (etc.)

In set-theoretical translations, then, individual constants and predicates or relation symbols are replaced by set terms which denote their extensions, and the element relationship $\in$ is inserted in between. This procedure is iterable for higher-order predications. Set-theoretical formalizations may be less intuitive than logical formalizations, but they have the advantage to be *type free*, that is, they allow for the translation of all sentences of (higher-order) type theory into the language of first-order PredL. For this reason, structuralist philosophers of science have argued that set-theoretical formalizations are preferable to logical formalizations (see [sec. 1.3.6](#)).

An extensional interpretation (D, I) of a language $\mathcal{L}$ is also called a *model* for $\mathcal{L}$. It plays the role of a “possible world” as expressible by means of $\mathcal{L}$. Often the models (D, I) of a sentence are written in the form $(D; a_1, \dots, a_n; R_1, \dots, R_m; f_1, \dots, f_k)$, in which instead of I , the designata of all individual constants, relational and function symbols assigned by I are listed. In this form, models are used by proponents of the (German) *structuralist* view and the (Anglo-Saxon) *semantic* view of theories. Here, a theory is represented by the set of its models, $M(T)$ – a method that goes back to Suppes ([1999](#), 249). By contrast, in the traditional “statement” view, a theory is represented by the set of its axioms, $Ax(T)$, together with semantic characterizations of its concepts.

The (traditional) “statement view” and the (structuralist) “non-statement view” have been set in opposition by Stegmüller ([1976](#)). This “opposition” is largely unjustified, for two reasons. *First*, the statement view is not (as is sometimes claimed) only a purely syntactical, but also a semantical view,

inasmuch as the statements of $\mathcal{L}$ are taken as semantically interpreted propositions (cf. [French 2008](#), 272). Since his turn in 1947, Carnap understood a linguistic framework always as a language $\mathcal{L}$ together with a semantic interpretation of $\mathcal{L}$ ([Carnap 1956b](#)). *Second*, model-theoretic theory reconstructions are not free from using statements (which is wrongly suggested by the name “non-statement view”). Rather, the class of models $M(T)$ of a theory T is defined by nothing but the set of its linguistic axioms $Ax(T)$, as follows: $M(T) =_{\text{def}} \{m \in M(\mathcal{L}) : m \models Ax(T)\}$, with $M(\mathcal{L})$ as the set of models for the language $\mathcal{L}$ in which T is formulated (cf. [Balzer et al. 1987](#), ch. 1). Therefore, there exist straightforward *translation relations* between linguistic and structuralist theory-representations. Here is a simple example:

(Prop. 3.1-2) *Linguistic and model-theoretic theory representation* (simplest case):

Linguistically by T's axioms:	Model-theoretically by T's models:
Semantic ‘axioms’:	Models: $M(T)$ = the class of models
$A_1 : I(F) \subseteq D. A_2 : I(G) \subseteq D.$	(D, F, G) such that $(D, F, G) \models \{A_1, A_2, A_3\}$.
Factual axiom $A_3 : \forall x(Fx \rightarrow Gx).$	

Often translation relations are more complicated, since structuralist theories contain several ingredients; for details see Andreas ([2010](#)) and Schurz ([2014b](#)).

3.1.3 The Content Type of Concepts

This classification is particularly significant in the philosophy of science. It applies only to *nonlogical concepts*. [Fig. 3.1-2](#) shows an *overview*.

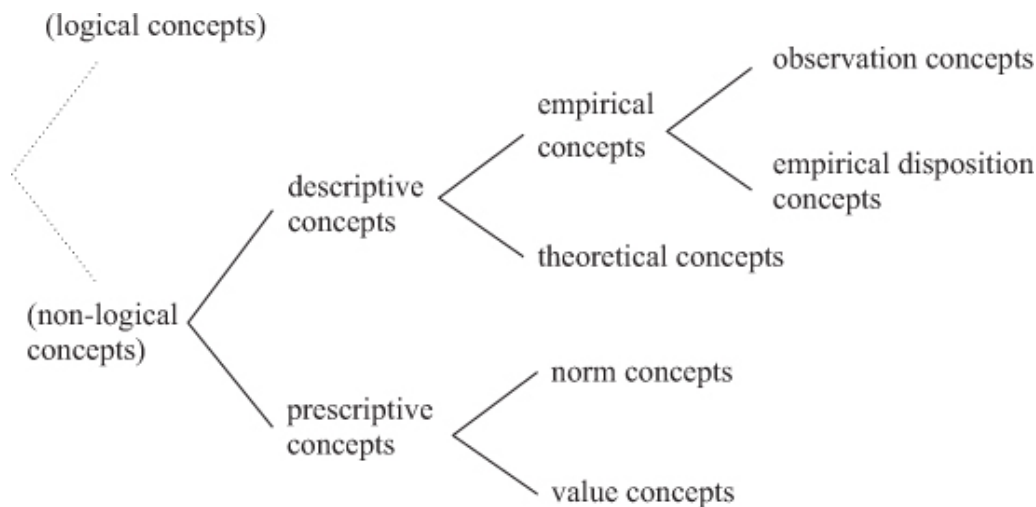


Figure 3.1-2 Classification of concepts according to their content type. [↩](#)

First, we have to distinguish between descriptive and prescriptive concepts. *Prescriptive* concepts consist of norm and value concepts ([Hare 1952](#)). There are only a few prescriptive concepts: the *normative* concepts of something being obligatory, forbidden, or permitted, and the qualitative, comparative, and quantitative *value* concepts (X is ethically or aesthetically valuable, or more valuable than Y , or valuable to a certain degree). Philosophy of science deals primarily with *descriptive* concepts and sentences which describe the facts of the world. There is an enormous number of (singular and general) descriptive concepts of various types.

The *most important* distinction within the descriptive concepts is the one between empirical and theoretical concepts, and within the empirical concepts the one between observation concepts and empirical dispositional concepts.

(1.) A general *observation concept in the narrow sense* (i.n.s.) expresses a property that is perceptible to the senses, or a complex of such properties. As we saw in [sec. 2.9.2](#), a property perceptible to the senses is one which can be *ostensively learned* by humans, independently of their background knowledge. *Examples*: “x is red,” “x is a raven,” “x is longer than y,” “the needle of the measuring instrument x is pointing to 5.” The classical and the early logical empiricists demanded that all descriptive concepts be observational concepts. In fact, only a *few* concepts are genuine observation concepts in this narrow sense.

Note the following subtlety: strictly speaking, a property as such is not perceptible, but only a certain *instantiation* of a property (a so-called *trope*) in a perceptible individual. “Being red,” for example, expresses an observational property as applied to visible objects; but in application to non-observable individuals such as microbes or molecules, the property of being red is not observable, as these individuals are too small to be visible. For this reason, Quine was right when he stressed that the primary level of “observability” is not that of concepts, but of “occasion sentences” like “this here is red” (1960, 35ff; 1992, 3). Putnam ([1962](#), 242f) saw in this fact a fundamental obstacle to a clear definition of observation concepts. Our criterion of ostensive learnability, however, solves this problem as follows: a primitive property is ostensively learnable, and thus observable i.n.s., iff it possesses at least *some* perceptible instantiations in perceptible individuals which may serve as training instances.

(2.) An *empirical dispositional concept* expresses an empirical disposition. By this, we understand the *lawlike regularity* of an individual to produce, under certain empirical circumstances or trigger conditions, a certain observable *result* that is also called the disposition’s “manifestation” ([Schrenk 2017](#), 45). Example: the empirical dispositional concept “x is soluble in water” is by definition equivalent to the following lawlike hypothesis about “x” which generalizes over time points: “whenever x is or were placed in water, it will or would dissolve.” Lawlikeness and counterfactual conditionals (“would”) go beyond the observable and presuppose inductive projectibility (see [sec. 3.6](#), 6.5.3). Nevertheless, dispositional concepts are *empirically operationalized* concepts, whose meaning is not dependent on a theory. Further examples: “x is hard,” “x weighs 4 kg on a pair of scales,” “person x is a German speaker.” Empirical dispositional concepts correspond to what the *operationalists* (e.g. [Bridgeman 1936](#)) demanded of *all* scientific concepts. In fact, though, theoretical concepts are not reducible to empirical dispositional concepts (see [sec. 3.10](#)).

(3.) An *empirical concept i.n.s.* is an observation concept or an empirical dispositional concept.

(4.) A *theoretical concept in the wide sense* (i.w.s.) is any descriptive concept which is not an empirical concept i.n.s., but still has a meaning. A theoretical concept i.w.s. does not have to be part of a scientific theory; it can also be based on speculation (e.g. “the prime cause of all being”).

(5.) A *theoretical concept i.n.s.* is a theoretical concept i.w.s., whose meaning is *introduced* by a scientific theory. If we want to make explicit the reference to a specific theory T, we speak of a *T-theoretical concept* ([Sneed 1971](#); refinements in [sec. 5.3.2](#)). *Examples*: physical force, kinds of molecules, biological genotype, psychological character, intention, and social order.

Convention: unless otherwise stated, we understand empirical concepts and theoretical concepts respectively in the narrow sense (i.n.s.).

(6.) As discussed in [sec. 2.9.1](#) (pro-5), a number of philosophers of science use the notion of an empirical concept in the wide sense (i.w.s.), meaning any concept that expresses a *measurable property* or ‘datum.’ Empirical concepts i.w.s. are theory dependent – their meaning depends on theories which describe the functioning of the measuring instruments and are presupposed to be unproblematic. The border between empirical concepts i.w.s. and theoretical concepts i.n.s. can only be drawn *pragmatically* and in a *context-dependent* way. By contrast, there exists an *objective* but gradual border between empirical concepts i.n.s. and theoretical concepts i.w.s.

3.1.4 The Gradation (Scale) Type of Concepts

Most properties in our world do not only occur or not occur; they come in certain *degrees*. The gradation or scale type of a concept reflects how finely the concept is graduated. Accordingly, scientific concepts may be classified as in the *overview* in [fig. 3.1-3](#).

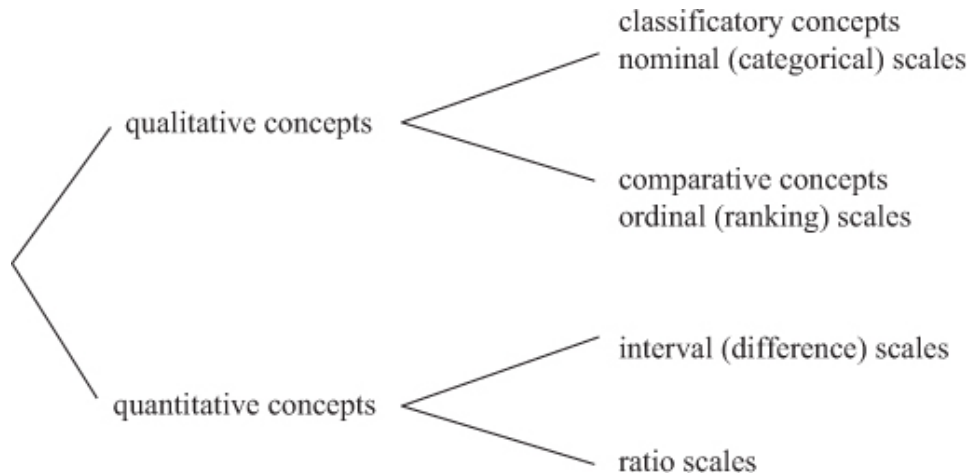


Figure 3.1-3 Classification of concepts according to their gradation type. [↗](#)

Qualitative concepts are not graduated, or more precisely, they are graduated only in the *yes-no* sense; comparative concepts are graduated in the sense of an *ordering*, and quantitative concepts are the most strongly graduated, that is, in a *continuous* way. In science, the graduation of a property is measured in terms of numbers, and we speak of the scale type of concepts (e.g. [Krantz et al. 2006](#), ch. 1; [Bortz 1985](#), 28–33). The *finer* graduated a concept is, the *sharper* the information we express with it, and the *stricter* the tests to which the hypotheses using it can be subjected. The types of gradation will now be discussed.

3.1.4.1 Qualitative-Classificatory Concepts and Nominal Scales

With the help of qualitative or classificatory concepts, we divide a domain D of objects into certain subgroups $D_1, \dots, D_n$, which are defined by their having certain properties or belonging to certain categories $C_1, \dots, C_n$. Here, it is demanded that: (a) the classification is *disjoint*, that is, the categories C_i must be mutually exclusive (the subgroups must not overlap), and (b) the classification is *exhaustive*, that is, the union of the subgroups must contain all of D .

The set $F =_{\text{def}} \{C_1, \dots, C_n\}$ of the categories of a classification is called a (classificatory) *concept family* ([Carnap 1971](#), 43ff). In the simplest case, an analytically independent concept Fx forms a two-valued or *binary* concept family $\{Fx, \neg Fx\}$: for example, x is either human or not. The conditions of disjunctivity and exhaustivity hold in this case purely logically. There are also concept families with more than two categories; for example, the family of color concepts $\{\text{red, yellow, green, and blue, } \dots\}$, or the family of social classes $\{\text{lower class, middle class, and upper class}\}$. The conditions of disjunctivity and exhaustivity are in these cases based on analytical conventions.

Every concept family F corresponds to a family concept or *attribute* X_F . For example, the concept family $\{\text{lower, middle, and upper class}\}$ corresponds to the attribute “social class of x .” Formally, X_F is a higher-order *function*, which assigns to every object in D its characteristic category, that is, $X_F : D \rightarrow \{C_1, \dots, C_n\}$. If we replace the categories C_i by *natural numbers* i , for example, “lower, middle, upper class” by “1,2,3,” then we obtain a nominal (or categorical) *scale*. For nominal scales, however, the use of numbering is inessential and arbitrary. In mathematics, a function X which when applied to various objects can take on various *numerical values* is also called a *variable*. Simple binary properties (F) correspond to a *binary* variable X_F with values F and $\neg F$. Observe that mathematical variables are functions; they should not be confused with variables in the logical sense.

3.1.4.2 Comparative concepts and ordinal scales

With the help of comparative concepts, the objects of a domain D are *ordered* according to the grade of a given property, that is, they are divided into *ranking groups* $D_1, \dots, D_n$ (see [fig. 3.1-4](#)): D_1 contains all D -objects with the highest grade of the given property, D_2 all D -objects with the next highest grade, etc. The

objects (“o”) within a rank group have the same grade. If the rank groups D_i are assigned natural numbers i , we have an *ordinal* or *ranking scale*. The corresponding assignment function $r: D \rightarrow \{1, \dots, n\}$ is an ordinally scaled variable, also called a ranking function.

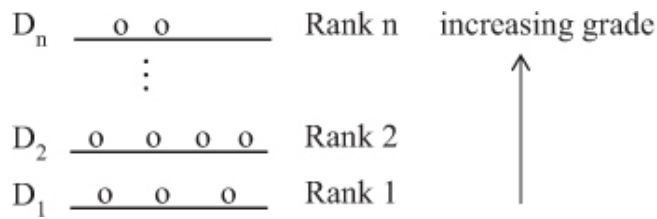


Figure 3.1-4 Ordinal scale of D . [↗](#)

The ranks of an ordinal scale give us only comparative information; quantitative differences between the ordinal ranks are unknown or undefined. In the *Mohs scale*, for example, the hardness of minerals is determined by a reciprocal *scratch comparison*: if mineral A scratches mineral B, but not vice versa, then A has a higher degree of hardness. Mohs used this method to devise a ten-point hardness scale: 1 (talcum), 2 (gypsum), 3 (calcium carbonate), up to 9 (corundum), and 10 (diamond). We must not conclude from this that the difference in hardness between talcum and gypsum is for example one-ninth of the difference between talcum and diamond. Similarly, the intensity of *psychological* attitudes can typically be measured only by an ordinal scale: for example, I can find one dish *tastier* than another without being able to say how much *tastier* it is.

An important method used in the empirical social sciences for *attitude surveys* consists in asking test persons (tps) to place their opinions on a given topic into a *preset* ordinal scale. A questionnaire on the attitudes of tps toward old age could for example contain the following question:

Please tick where appropriate: I believe that securing an old age pension is:

- ☐ very important ☐ quite important ☐ don't know
☐ not so important ☐ not important at all

The attitudes of the persons surveyed are thereby measured by an ordinal scale with five ranking levels, from not important at all (1) up to very important (5).

The basic logical concept of ordinal scales is that of a *smaller-or-equal* relation $\leq_P$ in regard to the given property P : $x \leq_P y$ states that the P -grade of x is less than or equal to that of y . With help of “ $\leq_P$,” one defines the relation of *P-equality* as $x \equiv_P y \leftrightarrow_{\text{def}} x \leq_P y \wedge y \leq_P x$ (x 's P -grade equals y 's), and the *strictly smaller* relation as $x <_P y \leftrightarrow_{\text{def}} x \leq_P y \wedge \neg y \leq_P x$ (x 's P -grade is smaller than y 's). The concepts $\leq_P$, $\equiv_P$, and $<_P$ are operationalized by *empirical* operations of comparison. In order to generate a *unique* ranking scale, these comparison operations must fulfill at least the axioms of a total *quasi-order*: *reflexivity* ($\forall x: x \leq_P x$), *transitivity* ($\forall x, y, z: x \leq_P y \wedge y \leq_P z \rightarrow x \leq_P z$), and *connexivity* ($\forall x, y: x \leq_P y \vee y \leq_P x$), that is, everything is comparable to everything else. Axioms and definitions entail the *anti-symmetry* of $\leq_P$ ($\forall x, y: x \leq_P y \wedge \neg x \equiv_P y \rightarrow \neg y \leq_P x$) and the fact that $\equiv_P$ is an *equivalence relation*, that is, fulfills the axioms of reflexivity, symmetry, and transitivity. (Note: an *order* in the mathematical sense arises from a quasi-order if no two D -objects occupy the same rank.)

The axioms of transitivity and connexivity are prerequisites for ordinal scaling. Whether they are at least nearly fulfilled is an *empirical* question (only reflexivity and anti-symmetry are fulfilled analytically). It is well known that the empirical preferences of tps are not always transitive (cf. [Garnham/Oakhill 1994](#), 184ff). In addition, all empirical *pair comparisons* have a certain *inaccuracy*. The relation of *approximate equality*, though, is no longer transitive when the grade differences between objects are very small. As an example, assume that a is 0.6 mm taller than b , b is 0.6 mm taller than c , but that only height differences of

more than 1 mm are observable. Then we observe that a is approximately equal to b, and b to c, but a is taller than c. If we place objects of this sort in a *preset* ordinal scale, we make their rank relations artificially transitive: the result may then depend on the order in which the objects are considered by the tp.

3.1.4.3 . Quantitative concepts and scales

These express the grade of a property P by means of a *numerical value* , along with a *unit of measurement* . Logically speaking, a quantitative concept is a function $f(x)$, which assigns to every object x in domain D a numeric value $r \in \text{Num}$ along with a unit of measurement e_f ; for example, $\text{mass}(x) = 5 \text{ kg}$, or $\text{temperature}(x) = 20^\circ \text{ Celsius}$. If $f(x)$ is a *discrete* magnitude, then f assigns integers to the objects; if $f(x)$ is a *continuous* magnitude, then f assigns real numbers to the objects. Almost all physical magnitudes are continuous; exceptions occur in quantum mechanics. Typical discrete magnitudes are counting numbers, for example, population sizes. In metaphysics, the nonspecified magnitude (e.g. “mass”) is also referred to as a determinable, and its numerical instantiation (e.g. “mass of 3.791 kg”) as a determinate property ([Wilson 2023](#)).

Directly observable are at best *comparative* magnitude relations: if we determine the length of a given distance by using a measuring tape, we can see whether this distance is longer, shorter, or approximately equal to certain scale divisions on the measuring tape. Converting a comparative concept into a quantitative concept by introducing an (empirical) measuring operation is called its *metrization* . In order to make a comparative property quantitatively measurable, certain empirical conditions must be satisfied, which are explicated axiomatically in *metrization theory* (see [Krantz et al. 2006](#)). On the other hand, any empirical measurement operation is based on certain *arbitrary conventions* , to which the results of respective measurements are relativized. Interval and ratio scales differ from each other in the amount of these arbitrary conventions. In an *interval scale* , both the *zero point* of the scale and the *unit of measurement* are set by convention. In a ratio scale, on the other hand, only the unit of measurement is conventional, while the zero point is empirically determined.

The empirical temperature scale, for example, is an interval scale. It is based on the fact that certain liquids, for example, mercury in a glass column, expand uniformly as the temperature increases. But there is no zero point of warmth or coldness which might be discerned through mere observation. In the *Celsius scale* ($^\circ \text{C}$), the freezing point of water is used as a conventional zero point, and the scale unit of 1°C is conventionally defined as one 100th of the length of expansion of mercury in a glass column as it rises from the freezing point to the boiling point of water. In the *Fahrenheit-scale* ($^\circ \text{F}$), the freezing point of water is set at 32°F , and the scale unit of 1°F corresponds to $5/9^\circ \text{C}$. So physically, it makes no sense to say that a liquid at 20°C is *twice as warm* as one at 10°C . This relationship depends on the artificial choice of the zero point; in the $^\circ \text{F}$ scale, the same relationship comes out as 68°F to 50°F . If, however, we have three liquids; a, b, and c at 10°C , 20°C , and 30°C , respectively, then it certainly does make sense in physical terms to say that the *temperature difference* between b and a is exactly as great as the one between c and b, and half as great as the one between a and c, as this is valid *independently* of the choice of zero point: in the difference between two temperature values, the zero point is canceled out. Because only statements about ratios of temperature intervals are objective, while statements about ratios of temperatures lack objective content, we talk of interval scales. Measurements of *position* and *time* also use interval scales, because the zero points of spatial and temporal coordinate systems are set by pure convention. We cannot say for example “the year 2000 is twice as late as the year 1000,” as this is relative to the birth of Christ as the starting point of the calendar. If our calendar began in the Middle Stone Age, we would have instead the years 12000 and 11000.

In *ratio scales* , unlike interval scales, the absolute zero point is empirically determined. For example, mass, volume, length (as opposed to spatial position), and duration (as opposed to temporal position) are ratio-scaled magnitudes. Physically, it makes sense to say that Peter, at 100kg, weighs twice as much as his 50 kg little brother. Because such ratio statements have objective content, we speak of ratio scales. Similarly, quantitative magnitudes in economics, such as amounts of money or goods, are based on ratio scales.

These magnitudes are also called *extensive* magnitudes, because when joining or “concatenating” two objects (a, b) into a whole (a o b), the resulting magnitude grows monotonically, that is, it holds that $f(a \text{ o } b) > f(a), f(b)$. ¹ Magnitudes such as temperature, for which this is not the case, are by contrast called *intensive* magnitudes – the temperature of say, water, is not increased if we increase the amount of water.

The metrization of extensive magnitudes leads to ratio scales. What is still arbitrary in ratio scales is the (conventional) choice of a *unit*. It makes no sense to say that the numerical value of Peter's weight is 100 – this is only valid if the unit of mass is assumed to be the kilogram; if we chose the gram as unit instead, the numerical value of Peter's mass would be 100,000. One subtle problem is the justification of the assumption that the chosen unit remains *constant over time* (see on this [Carnap 1966](#), ch. 8–9).

Only in simple *counting scales* is the unit naturally determined as one individual or item; such counting scales are also called *absolute scales*. The difference between various types of scales is mathematically described in terms of *admissible scale transformations*. The coarser a scale, the more scale transformations are admissible; for details see Krantz et al. ([2006](#), 11).

In theories of metrization, one proves, for example, the following: an extensive comparative magnitude relation $\leq_p$, for example, “is longer than,” over a domain of objects D is metrizable in the form of a ratio scale iff $\leq_p$ is a monotone quasi-order over D which satisfies the so-called *Archimedean* condition (see [Krantz et al. 2006](#), 73). The latter condition demands that every object b, no matter how large, can be balanced out by a sufficient number of copies of an arbitrary object a, no matter how small a is. For social preference relations, the Archimedean condition is often not satisfied – for example, one would be reluctant to say that the value of a human life might be balanced out by a sufficient number of banknotes (cf. [Schurz 1997](#), sec. 11.5.2).

Strictly speaking, the results of attitude measurements (e.g. 5 = very important, 3 = don't know, and 1 = unimportant) only have the level of an ordinal scale, or they lie somewhere “between” ordinal- and interval-scale levels. In spite of this, they are often evaluated using statistical tests which presuppose interval scales. To assess whether such an idealization is harmless in a given case, it is necessary to reevaluate the results of interval scale-based tests by an ordinal scale-based (or “non-parametric”) test (see [Bortz 1985](#), 31f, 178f).

So far, we have been concerned with *fundamental empirical* metrization. One speaks of *derived* metrization when the metrization of a concept is traced back to the metrization of other concepts, for example, density to the quotient of mass divided by volume. One speaks, finally, of *theory-dependent* metrization, when the metrization is based on a certain theory, for example, from physics ([Balzer 1985b](#), 6). In theory-dependent metrizations, a change in the scale level of an initially empirically defined metric may occur for theoretical reasons. An example of this is the *absolute* temperature scale after Kelvin ($^{\circ}\text{K}$), which postulates the absolute zero point of temperature at $-273.15^{\circ}\text{C} = 0^{\circ}\text{K}$. This assumption resulted from the *kinetic theory of gases*, according to which heat and temperature are to be attributed to the movement of molecules. In ideal gases, the volume (V) increases in a *linear* way with rising temperature ($V = k \cdot T$). If one extrapolates this relationship in a T×V-coordinate system to the left into the region of negative $^{\circ}\text{C}$ -temperatures, one reaches a temperature of -273.15°C , at which the volume of an ideal gas would have to be *zero*: here, the gas molecules would be crammed into an infinitely small space, and thus would be unable to move. Kelvin therefore identified this point as the absolute zero of temperature. His hypothesis that no cooling process could go below this absolute zero point has since been confirmed many times.

3.2 Classification of Sentences According to Their Content Type

Sentences may also be classified into several different dimensions. Most significant in philosophy of science is the classification of sentences according to their content type, which is treated of here, and according to their degree of generality, which is dealt with in sec 3.5. [Fig. 3.2-1](#) gives an *overview*.

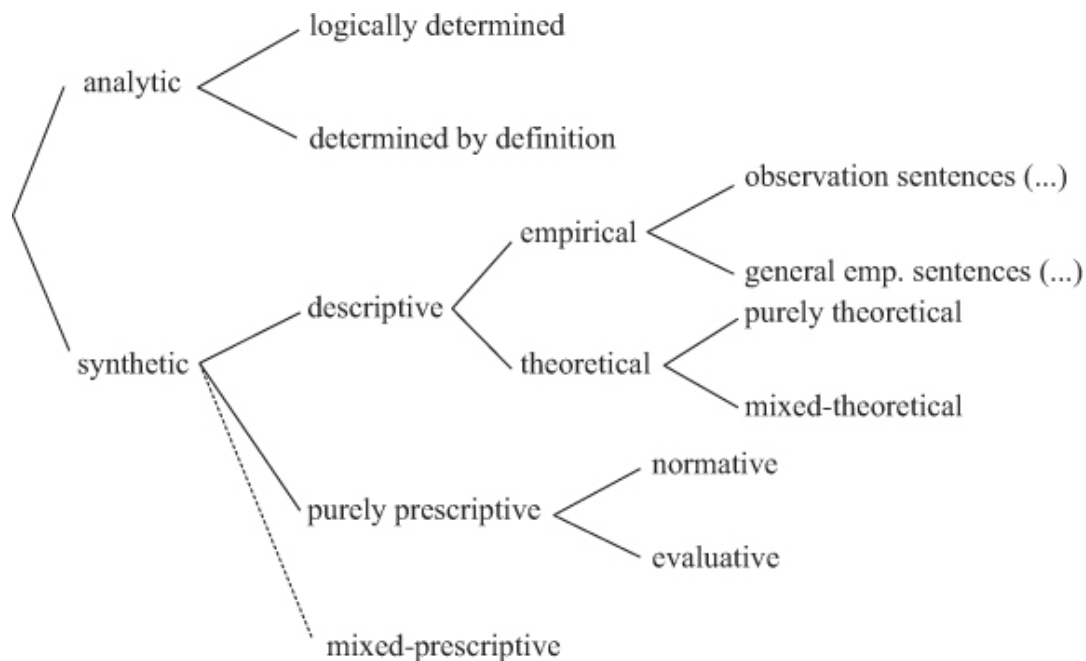


Figure 3.2-1 Classification of sentences according to their content type. [↗](#)

Analytically true sentences are those whose truth is determined by the laws of logic or by extralogical conventions of linguistic meaning, independently of the factual constitution of the world. Analytically false sentences are, by analogy, those whose falsity is determined in this way, and *analytic sentences* are those whose truth value (true or false) is determined in this way. Analytic sentences may be broken down further into logically determined sentences and sentences determined by definition or extralogical meaning convention ([sec. 3.3](#) – [3.4](#)).

By contrast, *synthetic* sentences are those sentences whose truth value is not analytically determined, but is dependent on the factual constitution of the world, or in the case of prescriptive-synthetic sentences, on our moral interpretation thereof. As with our content-based classification of concepts, we first divide synthetic sentences into descriptive and prescriptive ones; more precisely, into purely descriptive sentences, purely prescriptive sentences, and mixed sentences with descriptive and prescriptive components. Descriptive sentences tell us something about the factual composition of the world; prescriptive sentences on the other hand express propositions about norms or values. In the following subsections, we refine these distinctions.

3.2.1 Descriptive versus Prescriptive Sentences

Not every sentence containing a prescriptive concept is thereby prescriptive. Indeed, whenever a prescriptive concept lies within the *scope* of a *subjective attitude operator* – that is, a sentence operator expressing a subjective attitude toward a state of affairs – the prescriptive concept loses its prescriptive force. So for example, the sentences

Peter believes that stealing is wrong, – or
 Peter believes that you shouldn't steal,

are empirically descriptive, as they state something about Peter's factual attitude to values or norms. The same goes for all subjective attitude operators, like for example: person X believes, thinks, feels, says, wishes,..., that A. The significance of this fact for the value-neutrality thesis was worked out in [sec. 2.5.1](#) . We define:

(Def. 3.2-1) (a) A sentence is *descriptive* iff every prescriptive symbol occurring in it lies within the scope of a subjective attitude operator.

(b) A sentence is *purely prescriptive* iff all of its descriptive components lie within the scope of a prescriptive operator.

(c) A sentence which is neither descriptive nor purely prescriptive is called *mixed-prescriptive*. These sentences contain both descriptive components outside the scope of any prescriptive operator and purely prescriptive components outside the scope of any subjective attitude operator.

(3.2-1) *Examples of descriptive sentences* (logical structure at the right; prescriptive operators underlined):

- i. Peter steals. – Steals(Peter).
- ii. Peter believes that stealing is forbidden – Believes(Peter, Forbidden (stealing)).

Examples of purely prescriptive sentences:

- iii Stealing is forbidden. – Forbidden (stealing).
- iv If an action is under all circumstances morally bad, then it is forbidden. – $\forall a,c$ (Bad (a,c) $\rightarrow$ Forbidden (a)).

Examples of mixed-prescriptive sentences:

- v For persons suffering from hunger, stealing is allowed. – $\forall x$ (Suffering-hunger(x) $\rightarrow$ Permitted (Steals(x))).

Two remarks: (a) implicational connections between purely prescriptive sentences as in iv are themselves purely prescriptive. (b) Def. 3.2-1 is a *syntactic* definition; it can be refined by the method of *essential* sentence types in [def. 3.5-1](#).

3.2.2 Observation and Basic Sentences, Empirical and Theoretical Sentences

Building on the content-based classification of concepts in [sec. 3.1.3](#), we classify descriptive sentences according to their content type. Here, we have to address two subtleties which are rarely discussed in the literature. They have to do with the fact mentioned in [sec. 3.1.3](#), that instantiations of observable properties are only observable if the individuals possessing them are also observable.

First subtlety: an *empirical* sentence is usually defined as a sentence which (apart from logical concepts) contains only empirical concepts (cf. [Tuomela 1973](#), 23f; [Kutschera 1972](#), 260, 292). But this is insufficient. Consider a universal sentence “All Fs are Gs,” which contains only the observation concepts F and G. This universal sentence can only be called empirical if it refers solely to perceptible individuals. Whether this is the case depends on the concept F: only if it follows from the meaning of F that F can be applied *only* to perceptible individuals is the universal sentence an empirical sentence. For example, “all ravens are black” is a universal empirical sentence, because it follows from the meaning of “raven” that only perceptible individuals can be ravens. However, “all loose objects on Earth fall downwards” is not a universal empirical sentence if by “objects” we are also talking about microscopically small particles. The same applies to existential sentences. Lewis ([1970](#), 79) and van Fraassen ([1980](#), 13f) concluded from this problem that one cannot ascertain the empirical versus theoretical nature of a sentence from the meaning of its concepts. But as we will show now, this is indeed possible. For this purpose, we introduce the following *definition*: a quantifier in a given sentence has an *empirical range* iff this sentence is analytically equivalent with its quantificational restriction to observable individuals only. For a universal sentence $\forall xA$, this restriction is defined as the sentence $\forall x(Ox \rightarrow A)$, and for an existential sentence $\exists xA$ as $\exists x(Ox \wedge A)$, where “Ox” stands for “x is observable.” Otherwise, the quantifier has a *theoretical range*.

Second subtlety: observation sentences talk only of a finite or manageable number of individuals, because what humans can observe is of finite and limited complexity. It would be inadequate to limit observation sentences to *singular sentences* – that is, sentences *without* quantifiers – which contain only observation concepts. *Localized existential sentences* of the form “at point k in space-time, there is such-and-such an observable thing” are also observation sentences; for example, “in this basket (now) there is a red apple” ([Popper 1935](#) /2002, sec. V.28, speaks of “singular existential sentences”). This is the case because the

existence quantifier in such a sentence has empirical range, and because the limited segment of space-time is small enough to afford us a visual overview. For similar reasons, even *localized universal sentences* have to be counted as observation sentences, which have the following form: all (observable) individuals in this limited area of space-time have this-or-that observable property”; for example, “All apples in this basket are red.” Once again, the universal quantifier must have an empirical range; in addition, it is important that the individuals satisfying the antecedent predicate have a perceptible spatial extension that does *not overlap* with other individuals. Given these conditions, it follows from the meaning of “apple” and “basket” that there is only room for a finite and manageable number of apples in the described basket, so that we can determine the truth value of the sentence by observation.

This brings us to the following definition:

(Def. 3.2-2) If S is a synthetic-descriptive sentence, the following holds:

- a. S is an *observation sentence* iff apart from logical concepts, S contains only observation concepts, and S is either a singular sentence, or a *localized* existential or universal sentence, whose quantifiers possess empirical range.
- b. S is an *empirical sentence* (i.n.s.) iff apart from logical-mathematical concepts, S contains only empirical concepts i.n.s., and its quantifiers possess empirical range.
- c. S is a *theoretical sentence* i.w.s. iff S contains theoretical concepts i.w.s. or its quantifiers possess theoretical range.
- d. S is a *T-theoretical* sentence (or theoretical i.n.s.) iff S is a theoretical sentence i.w.s. and the theoretical concepts of S are introduced by the scientific theory T (refinements in [sec. 5.3.2](#)).

In analogy to [sec. 3.1.3](#) , we speak of an empirical sentence i.w.s. if apart from logical-mathematical concepts, it contains only empirical concepts i.w.s. (and quantifiers with an empirical range i.w.s.). Empirical sentences i.w.s. are dependent on unproblematic pre-theories and reflect what scientists pragmatically call ‘data.’ *Convention:* unless otherwise stated, we understand observation, empirical, and theoretical sentences i.n.s.

The class of empirical sentences is much more *comprehensive* than the class of observation sentences. Apart from observation sentences, it also comprises: (i) *universal* empirical sentences like for instance “all ravens are black,” (ii) singular empirical *disposition* sentences like “this is soluble in water,” and (iii) general empirical disposition sentences like “all sugar is soluble in water.” An important subgroup of empirical disposition sentences are empirical *measurement* sentences like “this tree is (rounded off) 10.52 meters high,” or “fully grown fir trees are more than 10 meters high.”

Theoretical hypotheses need not always be universal sentences. Singular or existence sentences can also be theoretical. For example, the singular sentence “this mineral is radioactive” expresses a theoretical hypothesis which can only be tested by means of theory-led measurements using complicated instruments. Finally, *theories* are *systems* of theoretical sentences (ch. 5). We stress here that *theories* or theoretical sentences may *also* contain empirical concepts, and even *must* do so in order to have empirical content, because some authors (e.g. [Putnam 1962](#) ; [Chalmers 1986](#) , 36; [Quine 1992](#) , 6–8) give the misleading impression that a concept, once it occurs within a theory, becomes a theoretical concept. According to our definitions, this is clearly not the case (recall [sec. 2.9.1](#) , contra-7).

Narrower still than the category of observation sentences is the category of *basic sentences* : basic sentences in the sense of Carnap ([1950b](#) , 67) and Hempel ([1965](#) , 277) are empirical atomic sentences or their negations, for example, this tree is (not) green. Basic sentences in the sense of Popper ([1935](#) /2002, ch. 5, sec. 28) are localized existential sentences, whose matrix is a conjunction of basic formulas, for example, “there is a black raven there.” All basic sentences are observation sentences, but not vice versa. In the Vienna Circle, it was hoped that basic sentences would serve to demarcate the *immediately* observable from that which can only be observed in a *mediated* way, via deduction from that which is immediately observable. For example, disjunctions such as “there is a crow or a raven sitting there” are not immediately observable, but they can be deduced from immediate observation sentences. Closer analysis, however, shows that this demarcation is problematic.

3.3 Logical Truth and Deductive Logic

According to philosophical tradition, logical truths are understood to mean “necessities of thought.” Both the history of philosophy and cognitive psychology show, however, that our *intuitive* ideas about necessities of thought are unreliable. Modern logic replaces the intuitive concept of necessity with exact concepts. The distinction introduced in [sec. 3.1.1](#) between logical and nonlogical concepts is of fundamental importance for the modern definition of logical truth. We will now present a general characterization of logical truth, which is applicable to any system of logic.

([3.3-1](#)) *Logical truth – preliminary characterization*: a sentence is logically true iff its truth follows logically from its syntactic structure and the meaning of its logical concepts alone.

In an equivalent reformulation of this idea, one can say: a sentence is logically true iff its truth follows from its *logical form* (this characterization goes back to [Wittgenstein’s Tractatus of 1921](#)). We find the logical form of a sentence by replacing all of its nonlogical concepts with metalanguage *variables* . The logical form of a sentence is what we get after the *formalization* of the sentence ([sec. 3.1.2](#)), but without any natural language interpretation of its nonlogical terms.

The characterization ([3.3-1](#)) is only preliminary, as the words “follows logically” in the defining phrase make it circular. The essential thing about a logical sentence form is that its nonlogical symbols may be *interpreted* in arbitrary ways. If every such interpretation makes the sentence true, then the sentence is logically true. We can therefore conclusively define the logical truth of a sentence, building on the idea in ([3.3-1](#)), as follows:

(Def. 3.3-1) A sentence is *logically true* (for short: *L-true*) iff its (logical) sentence form is true for every possible interpretation of its nonlogical symbols. We then also call this sentence form logically true.

Definition ([3.3-1](#)) goes back to Tarski ([1936b](#)). It is easily applicable to specific systems of logic, by specifying the permissible interpretations. In PropL, these are the assignments of truth values to atomic sentences. In PredL, these are the interpretations (D,I) explained in [sec. 3.1.2](#) . Note that in possible extensional interpretations (D,I), not only the interpretation function $I(\phi)$ for nonlogical symbols ϕ may vary, but also the individual domain D. Here are some examples:

(3.3-2) *Example of an L-true sentence*:

If everybody is mortal, then nobody is immortal.

Sentence form: if all Fs are Gs, then there exists no F which is not a G.

Formalization: $\forall x(Fx \rightarrow Gx) \rightarrow \neg \exists x(Fx \wedge \neg Gx)$

Further examples: $A \vee \neg A, A \rightarrow (A \vee B), \forall x Fx \rightarrow Fa, Fa \rightarrow \exists x Fx$.

The sentence form ([3.3-2](#)) is L-true, because it is true for every (extensional) interpretation of “F” and “G.” In other words, it is true in every possible world, as extensional interpretations correspond to possible worlds.

In analogy to [def. 3.3-1](#) , one defines a sentence as *logically false* (L-false) iff its sentence form is false for every possible interpretation of its nonlogical symbols. Examples of L-false sentence forms are $p \wedge \neg p$ (e.g. it will rain today and it will not rain), or $p \leftrightarrow \neg p$.

L-false sentences are also called *inconsistent* , or *contradictory* , because one may derive contradictions from them of the form “p and not p.” Sentences which are not L-false are called *consistent* . Consistency is also defined for *sets* of sentences: a sentence set is called consistent iff at least one interpretation verifies all sentences in the set; otherwise, it is called inconsistent (or contradictory). Finally, a sentence which is neither L-true nor L-false is called logically *contingent* . Examples of logically contingent sentences are Fa (e.g. Peter is ill) or $\forall x(Fx \rightarrow Gx)$ (e.g. all ravens are black). Most sentences are contingent.

With the help of the same method, we define the logical *validity* of arguments or inferences, which consists in the fact that the *truth* is transferred from the premises to the conclusion with *certainty* , that is, in all possible interpretations or ‘worlds’:

(Def. 3.3-2) An inference is *logically valid* (for short: valid) iff for the logical *form* of this inference the following holds: every possible interpretation which makes all premises true, also makes the conclusion true. – *Equivalently expressed*: ... iff there is no interpretation which makes all the premises true, but the conclusion false. – If an inference is valid, then we say that its conclusion *follows* (logically) from its premises.

(3.3-3) *Example of a logically valid inference*: Inference form:

Premise 1: All humans are mortal. $\forall x(Hx \rightarrow Mx)$

Premise 2: You are a human. Ha

Conclusion: Therefore, you are mortal. Ma

As with logical truth, logical validity is a matter of *logical form*. This fact implies that the validity of an inference does *not* depend on the *truth* of its premises (a common misunderstanding). The fact that an inference is valid only means: *if* its premises *were true*, then its conclusion *would* also have to be true. An inference with false premises can also be valid, like for example: “All humans are rich. You are a human. Therefore you are rich.” This inference is logically valid, since in all possible worlds in which all humans were rich, you being a human would be rich, too. That in the actual world not all humans are rich is merely a contingent fact. Of course, in order to apply an inference for purposes of *justification* or *prediction*, one must have good reasons for believing that its premises are true. Valid inferences, however, may also be used for purposes of *falsification*: in such cases, one infers from the falsity of the conclusion that *at least one* of the premises must be false.

To demonstrate that an inference form is semantically *invalid*, it suffices (according to [def. 3.3-2](#)) to find an interpretation that makes all the premises true but the conclusion false. On the other hand, semantic proofs of the *validity* in predicate logic are complicated and abstract. For this reason, logic also uses another procedure: the method of *deductive proofs*. With this method, one assumes some fundamental L-truths as *axioms* and some L-valid inferences as *rules* and aims to derive all other L-true sentences or valid inferences through the iterative application of axioms and rules in the form of “proofs” (see, e.g., [Klenk 1989](#), [Copi et al. 2010](#)). A full logical system consists of a language, a logical semantics, and a deductive proof system (or calculus).

3.4 Conventions of Meaning and Definitional Truths

As with L-true sentences, definitions such as “A bachelor is a hitherto unmarried man” are *analytical* truths, that is, their truth follows from the meaning of their concepts alone, independently of the factual constitution of the world. Differently from L-true sentences, however, the truth of definitions is determined not by the meaning of their logical concepts, but by conventions of meaning for their specific *nonlogical* concepts, in our example “bachelor,” “man,” and “hitherto unmarried.” The truth of sentences which are definitionally true is therefore *not* a matter of their logical form, but it is a matter of the conventions of meaning in the given language. So we define:

(Def. 3.4-1) (a) A sentence is *definitionally true* (or: *extralogically analytically true*) iff its truth is based on certain *conventions of meaning* for its *nonlogical* concepts.

In other words: a sentence is definitionally true iff it follows logically from the set of these conventions of meaning, without being itself logically true.

(b) A sentence is *analytically true* iff it is logically or definitionally true.

(3.4-1) *Examples of definitionally true sentences*:

a. All bachelors are male. $\forall x(Bx \rightarrow Mx)$

b. The length of the standard measure in Paris is one meter. $\text{Length}(a) = (1, m)$

(3.4-2) *Examples of synthetically true sentences:*

- a. All polar bears are white. $\forall x(Px \rightarrow Wx)$
- b. The length of this rod is one meter. $\text{Length}(b) = (1, m)$

The difference between analytic and synthetic truths can be illustrated as follows: if a zoologist sees the truth of (3.4-2)(a) as empirically confirmed, then his justification makes sense. But if someone reported to us that in spite of extensive empirical investigation, he has never met a female bachelor, then this person has clearly not understood the meaning of “bachelor”: bachelors are by definition (unmarried) men; no empirical research is necessary to know this. The examples also make it clear that definitional truths cannot be distinguished from synthetic truths on the basis of logical form, as (3.4-1)(a) and (3.4-2)(a) or (3.4-1)(b) and (3.4-2)(b) have the same logical form, respectively.

Although definitions cannot be *empirically* true or false, they can be more or less *empirically adequate* (see also [Hempel 1965](#), [sec. 6.5](#)). Before Linné, for example, animals and plants were classified by their phenomenological characteristics. Classification by common evolutionary descent, founded by Darwin a century later, turned out to be substantially more adequate, because it reflected a much larger number of common nomological characteristics. Nevertheless, nobody can be forbidden to define “fish” as “water-dwelling vertebrates.” The fact that dolphins, etc., would then also count as fish, although they descended from land-dwelling mammals, with whom they share many more common features than with the other (gill-breathing) fishes, does not disprove this definition – it does however, highlight its empirical inadequacy.

Some authors call particularly natural or empirically adequate definitions ‘true’ or ‘real definitions,’ as distinct from “nominal definitions” in the sense of linguistic conventions. This view goes back to Aristotle and presupposes an *essentialism*, according to which the world holds in itself a ‘true’ partition into natural kinds, which are characterized by certain necessary or “essential” properties (cf. [Fine 1994](#); [Schrenk 2017](#), [sec. 6.3](#)). What one counts as “essential” is however often unclear: there is ongoing debate about the optimal classification of biological taxa. Purely evolutionary classifications are generally held as superior to Linnaean classifications, but they would have to classify crocodiles not together with reptiles, but with birds, which is regarded as ‘unnatural’ by most biological taxonomists (cf. [Ridley 1993](#), 369). This shows that optimal definitions are not only dictated by the facts of the world, but also depend on epistemic matters. For this reason, we do not adopt the notion of ‘real definitions,’ because the demarcation between definitions and synthetic sentences would thereby become *unclear*. By “definitions,” we always understand more or less adequate semantic conventions.

There are three different kinds of sentences which are definitionally true, namely:

(D1) *Explicit definitions*: In these, a concept (the definiendum) is equated by convention with a complex of other concepts (the definiens). For general concepts, an explicit definition has the form of an *equivalence* (see [Suppes 1999](#), sec. 8.3):

(3.4-3)

(Definiendum) $\forall x: x \text{ is a bachelor}$	$\leftrightarrow_{\text{def}}$	(Definiens) $x \text{ is a hitherto unmarried man}$

Example: $\forall x: x \text{ is a bachelor} \leftrightarrow_{\text{def}} x \text{ is a hitherto unmarried man}$

Logical form: $\forall x(Dx_1 \dots x_n \leftrightarrow A)$, where in A only the individual variables $x_1 \dots x_n$ are free.

For singular concepts and function concepts, explicit definitions have the form of an identity, for example, $a =_{\text{def}} f(b_1, \dots, b_n)$, or $\forall x(f(x) =_{\text{def}} a \cdot x^n + b)$.

(D2) *Meaning postulates* (also called ‘*implicit*’ definitions) : These are semantic conventions, which like definitions play the role of extralogical-analytic *axioms*, that is, are not arrived at by logical derivation. These meaning postulates, however, do not have the form of an explicit definition, but some other form, mostly that of an implication. *Example*:

$\forall x(x \text{ is round} \rightarrow x \text{ has no corners})$.

(D3) *Derived definitorially true sentences*: These sentences are not logically true, but follow logically from sentences of type D1 or D2.

Example:

$\forall x(x \text{ is not male} \rightarrow x \text{ is not a bachelor})$.

Proper definitions must fulfill the criterion of *eliminability* of defined concepts (cf. Suppes 1957, 154). In order to satisfy this criterion, they have to satisfy the *noncircularity* requirement. For a single explicit definition, this criterion demands that the definiendum concept must not occur in the definiens. If this is the case, then the definiendum of explicit definitions of the form (3.4-3) can be eliminated in every sentence S in which it occurs, simply by replacing it by its definiens: the result is a sentence S* which is definitionally equivalent with S, that is, the equivalence $S \leftrightarrow S^*$ follows from the given definition.² In distinction to noncircular explicit definitions, meaning postulates do *not* fully characterize a concept and do not allow for its elimination.

For an entire *set* of definitions $\{D_1, \dots, D_n\}$, it must be required that this set can be arranged in the form of a *noncircular chain* of definitions, which means that every definiens in this chain contains only concepts which are either undefined or are the definiendum of an *earlier* definition in the chain (cf. von Savigny 1976, 116). An example of a noncircular chain of definitions is $Ax \leftrightarrow_{\text{def}} Bx \wedge \neg Cx$, $Bx \leftrightarrow_{\text{def}} Dx \wedge Ex$, $Dx \leftrightarrow_{\text{def}} \neg Fx$. If we carry out successive replacement in the definiens of Ax , we arrive at $Ax \leftrightarrow_{\text{def}} \neg Fx \wedge Ex \wedge \neg Cx$ as the basic definition of Ax , which contains only primitive (undefined) concepts in its definiens. An example of a circular chain of definitions is $\text{Necessary}(A) \leftrightarrow_{\text{def}} \neg \text{Possible}(\neg A)$, $\text{Possible}(A) \leftrightarrow_{\text{def}} \neg \text{Necessary}(\neg A)$, as after replacement, we arrive at the circular basic definition $\text{Necessary}(A) \leftrightarrow_{\text{def}} \neg \neg \text{Necessary}(\neg A)$.

A further category of definition is *recursive* definitions. They play an important role in logic and computer science. Without going into details, we only mention that in spite of their iterative structure, recursive definitions also satisfy the noncircularity requirement, since they are provably equivalent with (quite complicated) explicit definitions (cf. van Dalen et al. 1978, 157).

The central requirement for analytically true sentences is their empirical non-creativity, which is explicated as follows:

(Prop. 3.4-1) *Requirement of empirical non-creativity*: Analytically true sentences must neither have empirical content, nor create new empirical content in combination with the given system S of accepted background beliefs.

More precisely: A set of sentences Δ can *only* be analytically true if (a) the empirical content of Δ (in the sense of def. 3.7-1) is empty, and (b) the empirical content of S-united-with- Δ is the same as that of S.

Both requirements (a) and (b) must be fulfilled, in order to be able to say that the truth of such sentences is based on analytical convention. In Suppes' criterion (1999, 154) requirement (a) is missing, because he deals only with non-creativity *relative* to S.

Prop. 3.4-1 implies that *no concept may be defined twice* in different ways. It is well known that in the history of measurement, different conventions were suggested. Here are two examples:

- 1 meter = the length of the Platinum-Iridium-bar in Paris.
2. 1 meter = the length of a pendulum at sea level with an oscillation frequency of one second.

(1) + (2) however, implies the following *empirical consequence*: the length of the Platinum-Iridium-bar in Paris equals the length of a pendulum at sea level with an oscillation frequency of one second. We cannot regard (1) and (2) *simultaneously* as the definition of one meter. If (1) is the accepted definition, then (2) is an empirical hypothesis, and vice versa. Many "definitions" to be found in textbooks, for example, "Water = H_2O ," or "Child = person less than 12 years old," do not fulfill the requirement of empirical non-creativity. This is because the concepts being defined ("water," "child") have already been analytically introduced to common-sense background knowledge in other ways, for example, by ostensive characterizations. Such "definitions" should more rightly be called *explications* in the sense of Carnap (recall sec. 2.1) rather than

definitions. Also so-called “definitions” of theoretical concepts, for example, “force = mass times acceleration,” generate new empirical content and for this reason are not proper definitions (see [sec. 5.1](#)).

As long as a concept in a system of knowledge is characterized only by *one* explicit definition, this definition is provably always empirically non-creative (see [Shoenfield 1967](#), 57f). Also a *single* implicational meaning postulate for a new concept Fx such as $\forall x(Fx \rightarrow Ax)$ is not creative, because it is logically weaker than an explicit definition $\forall x(Fx \leftrightarrow Ax)$. Mostly though, an entire *system* of meaning postulates is introduced for one (or several) concepts; and for this, there is no guarantee of non-creativity. The designation of systems of meaning postulates as “implicit definitions” (which goes back to Hilbert) is therefore easily misunderstood, because “implicit definitions” behave quite differently from proper definitions.

3.5 Classification of Sentences According to Their Generality

The overview is shown in [fig. 3.5-1](#). The logical *strength* decreases successively with the transition from universal via singular to existential sentences: a universal sentence “Everything is an F” is logically stronger than its singular instance “this is an F” which is in turn logically stronger than the existential sentence “something is an F.”

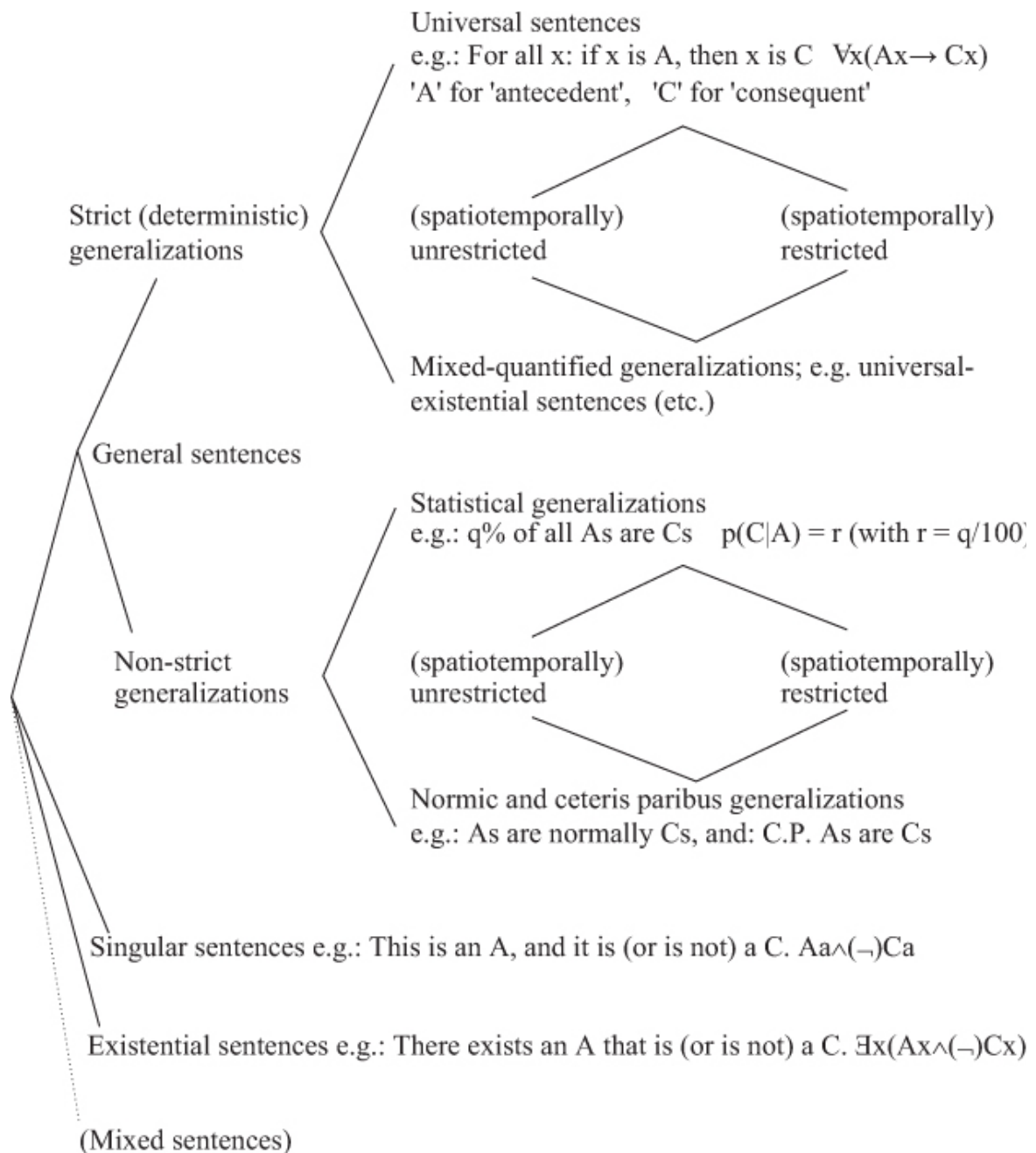


Figure 3.5-1 Classification of sentences according to their generality. ↩

1. *Strict generalizations* contain at least one universal quantifier. They include:

The *universal implication* “For all x: if Ax, then Cx” is the simplest form of a universal sentence. The if-component Ax is also called the *antecedent*, the then-component is called the *consequent*. In general, a *purely* universal sentence begins with one or more universal quantifiers, followed by a quantifier-free formula, which is called the *matrix* of the universal sentence and typically has the form of an implication, an equivalence, or a numerical identity. Universal sentences are also called *strict* or *deterministic* or *exceptionless* generalizations, since the truth of $\forall x(Ax \rightarrow Cx)$ implies for *every* individual a that if Aa obtains, the occurrence of Ca is inevitably determined.

Mixed-quantified generalizations contain both universal and existential quantifiers. *Example*: every action has a motive, formally $\forall x \exists y(\text{Action}(x) \rightarrow \text{Motive}(y,x))$. The meaning of generalizations with mixed

quantifiers depends on the *order* of quantifiers: $\exists x \forall y Rxy$ is a logically stronger statement than $\forall y \exists x Rxy$. *Example:* let “ Rxy ” stand for “ x is the cause of y .” Then $\forall y \exists x Rxy$ says that everything has some cause, but $\exists x \forall y Rxy$ says that some x is the cause of everything, that is, everything has the *same* cause, which is a much stronger statement. If the quantifiers are the same, though, their order is unimportant.

2. *Non-strict generalizations:* In this category, we include primarily *statistical* generalizations of the form “ $q\%$ of all A s are C s.” One calls this a statement of *conditional* probability and writes it down as $p(Cx|Ax) = r$, with $q = r/100$. The probability “ p ” is understood statistically, as a frequency or frequency limit ([sec. 3.9](#)). Although statistical generalizations are general sentences, they are *not* universal sentences. The statistical generalization “ $q\%$ of all A s are C s” says nothing about single individuals of type A , rather it only says something about the class of A s, that is, that $q\%$ of the individuals of the class of A s are located in the class of C s (see [fig. 3.5-2](#)). For this reason, there are no logical consequence relations, but only inductive-probabilistic relations between statistical generalizations and singular sentences, if the domain of individuals is unknown or potentially infinite.

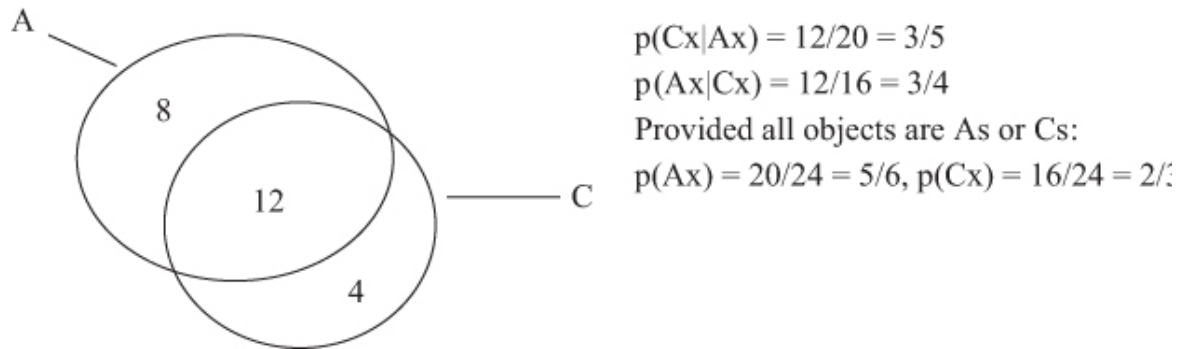


Figure 3.5-2 Statistical generalizations. ↩

A further kind of non-strict generalizations are *normic* generalizations or normal case hypotheses of the form “ A s are *normally* C s” – for example, birds can normally fly (the term “normic” goes back to [Scriven 1959a](#)). For reasons explained in [sec. 4.5.2](#) , we can assume that a normic law analytically *implies* a numerically unspecific statistical generalization of the form “*most* A s are B s,” without being semantically equivalent with it. A third kind of non-strict generalizations are *ceteris paribus* (CP) generalizations, which say that Ax leads to Bx as long as other interfering factors are either *held constant* (comparative CP generalizations), or are *absent* (exclusive CP generalizations, see [sec. 4.5.1](#)).

What statistical, normic, and CP generalizations have in common, is that they allow for *exceptions* – these are instances which verify the antecedent and falsify the consequent predicate. Therefore, inferences from them are described by a non-monotonic consequence relation in the sense of [sec. 2.7.4](#) . Strict laws forbid such exceptions and are falsified by them.

Both strict and non-strict generalizations can be unrestricted or restricted in space and time. They are restricted if the antecedent predicate contains a restriction to a finite area of space and time – for example, “all swans in *Europe* are white,” as opposed to “all swans are white.” Unrestricted universal sentences are also called *spatiotemporally universal* ([sec. 6.5.4](#)).

3. *Singular* are all sentences which contain *no quantifiers* , and also no statistical probability function. Singular sentences are composed of atomic sentences using propositional operators (Fa , $\neg Fa$, $Fa \wedge Rab$, etc.). They always say something about one or more individual things.

4. *Existential sentences:* Pure existential sentences begin with one or more existential quantifiers, followed by a quantifier-free matrix formula. Usually, the matrix has the form of a conjunction $\exists x(Ax \wedge \pm Cx)$; for example, there is a black swan. Such existential sentences may also be unrestricted or restricted, depending on whether Ax contains a restriction in space and time or not.

5. *Mixed sentences* , finally, are propositional combinations of the above sentence types, like for instance $\forall x(Fx \rightarrow Gx) \vee Ha$. They play a subordinate role.

The table in [prop. 3.5-1](#) summarizes the most important logico-deductive relationships between the listed kinds of sentences.

(Prop. 3.5-1) *Notation*: “ $X \Vdash Y$ ” stands for “ Y follows logically from X ” (“ $\dashv$ ” for “logically equivalent”).

1. Universal sentence $\Vdash$ singular sentence
All As are Cs $\Vdash$ if a is A, then a is C
2. Universal sentence and singular sentence $\dashv$ singular sentence (*explanation scheme*)
All As are Cs, and a is A $\dashv$ a is C
3. Singular sentence falsifies strictly universal sentence (*falsification scheme I*)
Singular sentence $\dashv$ negation of a universal sentence
a is A and not C $\dashv$ not all As are Cs
4. Existential sentence falsifies strictly universal sentence (*falsification scheme II*)
There is an A that is not a C $\dashv$ not all As are Cs
5. Universal sentence is logically equivalent to the negation of an existential sentence.
All As are Cs $\dashv$ $\dashv$ there is no A that is not a C
6. Singular sentence $\dashv$ Existential sentence
a is an A $\dashv$ there is an x that is an A
7. There are no logical consequence relations between non-strict generalizations and singular sentences, only statistical or epistemic relations of probability.

The above characterizations of kinds of sentences are of a *syntactic* nature. Syntactic characterizations are not always invariant with respect to (w.r.t.) logically equivalent (L-equivalent) transformations. For instance, the singular sentence Fa is L-equivalent to the universal sentence $\forall x(x = a \rightarrow Fx)$ (all x which are identical with a have property F). This sentence, however, is not a universal sentence in the *semantic* sense, as it only says something about the individual a . If characterizations of sentences should express genuine properties of the *propositions* thereby expressed, they must be invariant w.r.t. L-equivalent sentence transformation, since L-equivalent sentences express the same proposition. A standard method of solving this problem is the method of *essential* sentence types (see [Hempel/Oppenheim 1948](#) , 271). ³ Here, one regards a sentence as essentially universal (or existential) only if it is not L-equivalent to a singular sentence:

(Def. 3.5-1) *Essential sentence types (according to their generality)*:

1. A sentence S is an essentially universal/existential sentence iff S is a syntactically universal/existential sentence and not L-equivalent to a syntactically singular sentence.
2. By contrast, a sentence S is an *essentially* singular sentence iff S is L-equivalent to a syntactically singular sentence.

The division of kinds of sentences according to their generality is largely *independent* of the content-based classification of sentences given in [sec. 3.2](#) : the kinds of sentences listed above can be either empirical or theoretical sentences, depending on which nonlogical concepts occur in these sentences, and which range their quantifiers possess. Unrestricted universal or existential sentences can however never be observation sentences.

A further distinction, which runs *crosswise* to our distinctions up to now, is the one between *evidences* and *hypotheses* . It relates to the *epistemic status* of our knowledge. Evidence statements are very well-secured sentences – actual observation sentences are prototypical for this category. Hypotheses are less well-secured sentences, which are in need of further testing – prototypical for this are general sentences, or potential observation sentences (i.e. observable predictions).

PHILOSOPHY OF SCIENCE

A Unified Approach

GERHARD SCHURZ

Second Edition



PHILOSOPHY OF SCIENCE

A Unified Approach

GERHARD SCHURZ

Second Edition



“I consider this book to be a unique work both with regard to the sheer amount of material it covers but also with regard to the clarity, depth and insight with which these matters are dealt with by one of Europe’s foremost analytic philosophers. I would not hesitate to use it in an advanced course on the subject, and I believe it would work well also for an introductory course given the careful way in which the book is structured into introductory and advanced material.”

– **Erik J. Olsson** , *Lund University*

“ *Philosophy of Science: A Unified Approach* revamps and bolsters the logical empiricist way to think about and to practice philosophy of science. It combines technical precision and rigor with philosophical breadth and depth. It delves into all major philosophical problems and issues in general philosophy of science. It elegantly presents the received positions and approaches, but it also articulates novel views and arguments. It systematically explains the necessary technical concepts and makes them do deep philosophical work. *Philosophy of Science: A Unified Approach* lives up to its promise to highlight and defend the methodological unity of the sciences. Gerhard Schurz has written a superb introduction to philosophy of science; novices and experts alike have a lot to learn from it.”

Stathis Psillos , *University of Western Ontario and University of Athens*

“This is an excellent introduction to the general philosophy of science. It is lucidly written and very well structured in elementary and advanced subsections. Inspired by the classic authors of the 20th century, such as Carnap and Popper, and guided by the methodological version of the idea of the unity of the sciences and the humanities, Gerhard Schurz shows how it stands with the main topics in the first decades of the 21st century.”

– **Theo Kuipers** , *University of Groningen*

Philosophy of Science

Philosophy of Science: A Unified Approach, Second Edition, combines a general introduction to philosophy of science with an integrated survey of all its important subfields. As this book's subtitle suggests, this excellent overview is guided methodologically by "a unified approach" to philosophy of science. Behind the diversity of scientific fields, one can recognize a methodological unity. This unity is worked out in this book, revealing all the while important differences between subject areas.

This book offers a two-part approach that makes it an excellent introduction for readers new to the field and a useful resource for more advanced students. Each chapter is divided into two sections. The first section assumes no foreknowledge of the subject introduced, and the second section builds upon the first by bringing into the conversation more advanced, complementary topics from contemporary research.

Definitions, key propositions, examples, and figures complement all of the core material. At the end of every chapter, there are selected readings and exercise questions (with solutions and answers to many of them at the end of this book). This book also includes a comprehensive bibliography and an index.

The Second Edition has been revised to make the technical material much more accessible for students outside of the sciences and philosophy. It also incorporates new research over the last decade, including significantly expanded coverage of values in science, mechanistic explanations, and inductive risk. An extended index of subjects that can also be used as a glossary has been added to the back of this book.

Gerhard Schurz (MSc 1980, PhD 1983) is Professor of Philosophy at the Heinrich Heine University Düsseldorf (since March 2002) and member of the Leopoldina (German Academy of Sciences) and the International Academy for Philosophy of Science. Before 2002, he was an Associated Professor at the University of Salzburg and a Guest Professor at the University of California at Irvine and Yale University. His research areas include philosophy of science, epistemology, logic, and cognitive science. He is the author of 270 research articles and 12 books.

Philosophy of Science

A Unified Approach

Second Edition

Gerhard Schurz



Designed cover image: Getty Images

Second edition published 2026
by Routledge
605 Third Avenue, New York, NY 10158

and by Routledge
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2026 Gerhard Schurz

The right of Gerhard Schurz to be identified as author of this work has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

The Open Access version of this book, available at www.taylorfrancis.com, has been made available under a Creative Commons Attribution-ShareAlike 4.0 (CC BY-SA) International license.

Any third party material in this book is not included in the OA Creative Commons license, unless indicated otherwise in a credit line to the material. Please direct any permissions enquiries to the original rightsholder.

For Product Safety Concerns and Information please contact our EU representative
GPSR@taylorandfrancis.com. Taylor & Francis Verlag GmbH, Kaufingerstraße 24, 80331 München, Germany.

Trademark notice : Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

First edition published by Routledge 2013

Library of Congress Cataloging-in-Publication Data

Names: Schurz, Gerhard, 1956– author

Title: Philosophy of science : a unified approach / Gerhard Schurz.

Other titles: Einführung in die Wissenschaftstheorie. English

Description: 2nd edition. | New York, NY ; Abingdon, Oxon : Routledge, 2026. | Includes bibliographical references and index. | Text in English. Translation from German.

Identifiers: LCCN 2026005541 | ISBN 9781041145820 hardback | ISBN 9781041145813 paperback | ISBN 9781003675075 ebook

Subjects: LCSH: Science—Philosophy

Classification: LCC Q175 .S413414 2026

LC record available at <https://lcn.loc.gov/2026005541>

ISBN: 978-1-041-14582-0 (hbk)

ISBN: 978-1-041-14581-3 (pbk)

ISBN: 978-1-003-67507-5 (ebk)

DOI: [10.4324/9781003675075](https://doi.org/10.4324/9781003675075)

Typeset in Sabon
by codeMantra

Contents

[List of Figures, Definitions, and Key Propositions](#)

[List of Symbols](#)

[Foreword](#)

[Foreword to the Second Edition](#)

[1 Introduction: Where Do We Stand?](#)

- [1.1 Tasks and Aims of Philosophy of Science](#)
- [1.2 On the Historical Development of Philosophy of Science](#)
- [1.3 Philosophical Positions in Contemporary Philosophy of Science](#)
 - [1.3.1 Logical Empiricism](#)
 - [1.3.2 Critical Rationalism](#)
 - [1.3.3 The Historical Account and Relativism](#)
 - [1.3.4 Metaphysical and Scientific Realism](#)
 - [1.3.5 Naturalism and the Cognitive Turn](#)
 - [1.3.6 Structuralist and Semantic View of Theories](#)
 - [1.3.7 Pragmatic Philosophy of Science](#)
 - [1.3.8 Hermeneutics and Critical Theory](#)
 - [1.3.9 Radical Constructivism, Postmodernism, and the Science Wars](#)
- [1.4 Introductory Literature and Exercises](#)

[2 The Question of Unity: In Search for Common Foundations of the Sciences](#)

- [2.1 Normative or Descriptive? The Method of Rational Reconstruction](#)
- [2.2 Common Epistemological Assumptions](#)
- [2.3 Common Methodological Features](#)
- [2.4 Scientific Disciplines and Their Classification](#)
 - [2.4.1 The Special Status of Mathematical Sciences](#)
 - [2.4.2 Testing the Common Foundations of Science](#)
 - [2.4.3 Two Further Classifications of Sciences](#)
- [2.5 The Question of Value Neutrality](#)
 - [2.5.1 Max Weber's Value-Freedom Postulate](#)
 - [2.5.2 Value Neutrality and Means-End Inferences](#)
 - [2.5.3 Explication of the Value-Neutrality Requirement](#)
- [2.6 The Demarcation Problem and the Methodological Unity of Sciences](#)
 - [2.6.1 Demarcation](#)

2.6.2 Methodological Unity instead of Physical Reductionism

2.7 Scientific Inference and Argumentation

2.7.1 Deduction and Induction

2.7.2 Popper and the Significance of Inductive Inference in the Sciences

2.7.3 Abduction and Inference to the Best Explanation

2.7.4 Monotonic and Non-monotonic Inferences

Complementary and Advanced Topics

2.8 How Much Metaphysics is Presupposed by Science?

2.8.1 Truth, Idealism, and Realism

2.8.2 Constructive Realism versus Radical Constructivism

2.9 On the Theory (In)dependence of Observation

2.9.1 Theory-Dependency of Observation – Pros and Cons

2.9.2 A Definition of Theory-Neutral Observation Concepts

2.10 Value Neutrality in the Context of Contemporary Debates

2.10.1 The Role of Values in Science

2.10.2 The Argument of Inductive Risk

2.10.3 Meta-ethical Justification of Value Neutrality and the Difference between Experience and Valuation

2.11 The Justification of Induction: An Unsolvable Problem?

2.11.1 Is Induction Rationally Unjustifiable?

2.11.2 The Optimality of Meta-induction

2.12 Introductory Literature and Exercises

3 The Conceptual Toolkit: Language, Logic, and Probability

3.1 Kinds and Classification of Concepts

3.1.1 The Logical Type of Concepts

3.1.2 Syntax and Semantics, Statements, and Models

3.1.3 The Content Type of Concepts

3.1.4 The Gradation (Scale) Type of Concepts

3.2 Classification of Sentences According to Their Content Type

3.2.1 Descriptive versus Prescriptive Sentences

3.2.2 Observation and Basic Sentences, Empirical and Theoretical Sentences

3.3 Logical Truth and Deductive Logic

3.4 Conventions of Meaning and Definitional Truths

3.5 Classification of Sentences According to Their Generality

3.6 General Sentences, Lawlikeness, Determinism, and Indeterminism

3.7 Content of Sentences and Kinds of Content

3.8 Verification, Falsification, Confirmation, and Disconfirmation

3.9 Objective (Statistical) and Subjective (Epistemic) Probability

Complementary and Advanced Topics

3.10 Disposition Concepts

3.11 Challenges to the Logical-Definitional-Synthetic Distinction

- [3.12 Relevance and Irrelevance in Logical Inferences](#)
- [3.13 More on Probability](#)
 - [3.13.1 Mathematical Laws of Probability](#)
 - [3.13.2 Problems of Objective-Statistical Probability Concepts](#)
 - [3.13.3 Problems of Subjective Probability Concepts](#)
 - [3.13.4 Connections between Objective and Subjective Probabilities: Principal Principle and Narrowest Reference Class](#)
- [3.14 Introductory Literature and Exercises](#)

4 A Question of Fit: Law Hypotheses and Their Empirical Testing

- [4.1 The Condition of Relevance](#)
 - [4.1.1 Relevance of Strict Laws](#)
 - [4.1.2 Relevance of Statistical Laws](#)
- [4.2 Testing of Strict Law Hypotheses for Truth and Relevance](#)
 - [4.2.1 The Method of Agreement and Difference](#)
 - [4.2.2 Methodological Induction: How to Discover Strict Law Hypotheses](#)
- [4.3 Testing Statistical Laws](#)
 - [4.3.1 Testing for Likely Truth and Acceptance Intervals](#)
 - [4.3.2 Discovery of Statistical Hypotheses and Confidence Intervals](#)
 - [4.3.3 Test Statistics and Inference Statistics](#)
 - [4.3.4 Testing for Relevance and Significant Differences](#)
 - [4.3.5 Statistical Representativeness](#)
 - [4.3.6 Sources of Error in Statistical Methods](#)
 - [4.3.7 Applying Statistical Hypotheses to Individual Cases](#)
- [4.4 Correlation and Causality](#)
 - [4.4.1 Hidden Variables](#)
 - [4.4.2 Causal Direction](#)

Complementary and Advanced Topics

- [4.5 Ceteris Paribus Laws and Normic Laws](#)
 - [4.5.1 Comparative versus Exclusive CP Laws](#)
 - [4.5.2 Normic Laws and Their Evolution–Theoretic Foundation](#)
- [4.6 Probability Distributions for Continuous Variables](#)
- [4.7 From Classical Statistics to Bayesian Statistics](#)
 - [4.7.1 The Likelihood Intuition and Its Bayesianism Justification](#)
 - [4.7.2 Objective Bayesianism and the Principle of Indifference](#)
 - [4.7.3 Subjective Bayesianism and Convergence of Opinions](#)
- [4.8 Introductory Literature and Exercises](#)

5 Going beyond Experience: Theories and Their Empirical Evaluation

- [5.1 Theoretical Concepts and Multiple Laws of Correspondence](#)
- [5.2 The Example of Newtonian Physics](#)
- [5.3 The Structure of Scientific Theories](#)
 - [5.3.1 Components of Theories](#)
 - [5.3.2 Theory Nets, Intertheoretical Relations, and Kinds of Theoretical Concepts](#)

- 5.4 [Methodological Features of \(Good\) Scientific Theories](#)
 - 5.4.1 [System Character: Holism of Meaning, Content, and Theory Testing \(Duhem's Thesis\)](#)
 - 5.4.2 [Empirical Creativity, Global Unification, and Novel Predictions: Answers to Hempel's Dilemma](#)
- 5.5 [The Example of Piaget's Theory of Cognitive Development](#)
- 5.6 [The Dynamics of Scientific Theories](#)
 - 5.6.1 [Lakatos' Model of Theory Revision](#)
 - 5.6.2 [Theory Evaluation and Theory Progress](#)
- 5.7 [Instrumentalism and Realism: The Ontology of Scientific Theories](#)

[Complementary and Advanced Topics](#)

- 5.8 [Instrumentalism-Realism Controversies in Philosophy of Science](#)
 - 5.8.1 [The Ramsey Sentence of a Theory: An Instrumentalist Challenge](#)
 - 5.8.2 [The Realist Interpretation of the Ramsey Sentence](#)
 - 5.8.3 [The No-Miracles Argument](#)
 - 5.8.4 [The Pessimistic Meta-Induction](#)
 - 5.8.5 [The Empirical Underdetermination of Theories](#)
 - 5.8.6 [Intertheoretical Correspondence, Structural and Selective Realism](#)
 - 5.8.7 [Criteria for Empirical Significance](#)
 - 5.8.8 [Probabilistic Analysis of the NMA, Alternative Theories, and Local Realism](#)
- 5.9 [General Accounts of Confirmation](#)
 - 5.9.1 [Hypothetico-Deductive Confirmation](#)
 - 5.9.2 [Probabilistic Confirmation and Bayesianism](#)
 - 5.9.3 [Bayesian Pseudo-Confirmation through Content Cutting versus Genuine Confirmation](#)
 - 5.9.4 [Goodman's Paradox](#)
 - 5.9.5 [Confirmation and Acceptance](#)
- 5.10 [Testing Theories: Parameter Fitting and Use-Novel Evidence](#)
 - 5.10.1 [Post-Factum Speculations](#)
 - 5.10.2 [Curve Fitting](#)
- 5.11 [Non-Confirmational Accounts of Theory Evaluation](#)
 - 5.11.1 [Truthlikeness](#)
 - 5.11.2 [Unification, Coherence, and Simplicity](#)
- 5.12 [The Problem of Language Dependence](#)
- 5.13 [Introductory Literature and Exercises](#)

[6 In Search of Causes: Explanation and All That Goes with It](#)

- 6.1 [The Deductive-Nomological Model of Explanation](#)
- 6.2 [Explanation versus Prediction and Justification](#)
 - 6.2.1 [Predictions without Explanatory Function](#)
 - 6.2.2 [Causality and Lawlikeness](#)
 - 6.2.3 [Explanations without Predictive Function](#)
 - 6.2.4 [Logical Problems of Irrelevance and Redundancy](#)

6.3 Probabilistic Explanation Models

6.3.1 Inductive-Statistical Explanation

6.3.2 Maximal Specificity

6.3.3 Probabilistic and Causal Relevance Models

6.3.4 Conflicting Intuitions about the Magnitude of Explanatory Probabilities

6.4 Further Types of Explanation

6.4.1 Mechanistic Explanations

6.4.2 Normic Explanations and the Explanation of Human Actions

6.4.3 Explanation and Understanding as Unification

Complementary and Advanced Topics

6.5 Lawlikeness

6.5.1 Laws of Nature versus System Laws

6.5.2 Lawlikeness in the Wide versus Narrow Sense

6.5.3 Lawlikeness i.w.s. and Counterfactuals

6.5.4 Lawlikeness i.n.s., Spatiotemporal Universality, and Maxwell's Condition

6.5.5 The Best System Account

6.5.6 Physical Necessity and Independent Possibility Knowledge

6.6 Causality

6.6.1 Singular and General Causality

6.6.2 Regularity Accounts

6.6.3 Counterfactual Accounts

6.6.4 Causation and the Direction of Time

6.6.5 Causal Process Theories

6.6.6 Interventionist Accounts

6.6.7 Causality as a Theoretical Concept

6.6.8 The Theory of Causal Nets

6.6.9 Causal Discovery and Empirical Content of TCN

6.6.10 From Variables to Events: Overdetermination and Causal Preemption

6.7 Introductory Literature and Exercises

Bibliography

Solutions to Selected Exercises

Index of Names

Index of Subjects

List of Figures, Definitions, and Key Propositions

Figures

- [2.1-1 The method of rational reconstruction](#)
- [2.3-1 The three levels of the scientific method](#)
- [2.4-1 Classification of the factual sciences](#)
- [2.5-1 Schematic representation of the value neutrality requirement](#)
- [2.7-1 The inductive-deductive scheme](#)
- [2.7-2 Interaction of epistemic induction and abduction](#)
- [2.9-1 Duck-rabbit](#)
- [2.9-2 Mueller-Lyer illusion](#)
- [2.9-3 Expected learning curves in ostensive learning experiments](#)
- [3.1-1 Division of concepts according to their logical type](#)
- [3.1-2 Classification of concepts according to their content type](#)
- [3.1-3 Classification of concepts according to their gradation type](#)
- [3.1-4 Ordinal scale](#)
- [3.2-1 Classification of sentences according to their content type](#)
- [3.5-1 Classification of sentences according to their generality](#)
- [3.5-2 Statistical generalizations](#)
- [3.8-1 Testability of generalizations](#)
- [3.9-1 Non-monotonicity of conditional probabilities](#)
- [3.13-1 Three binomial distributions](#)
- [4.1-1 Irrelevance and joke](#)
- [4.3-1 Acceptance interval](#)
- [4.3-2 Acceptance and confidence interval](#)
- [4.3-3 Probability distribution of sample differences and significant difference](#)
- [4.4-1 Common cause](#)
- [4.4-2 Common cause vs. intermediate cause](#)
- [4.4-3 Spurious causal independence](#)
- [4.4-4 What is the direction of causality?](#)
- [4.6-1 Normal distribution](#)
- [4.7-1 Shortest interval](#)
- [4.7-2 Transformation of uniform into non-uniform density](#)
- [5.1-1 One theoretical feature causes many empirical dispositions](#)
- [5.1-2 Beam balances and spring scales on Earth and Uranus](#)
- [5.2-1 Theory lattice in classical mechanics](#)

[5.3-1,2,3 Subdivisions of a theory](#)
[5.3-4 Theory net of physical and chemical theories](#)
[5.6-1 Historical theory sequence](#)
[5.6-2 Theory comparison](#)
[5.6-3 Empirically complementary theories](#)
[5.6-4 Theories of optics after Bunge](#)
[5.10-1 Curve fitting and use-novel evidence](#)
[6.6-1 Explanation of screening-off](#)
[6.6-2 Three statistically strongly indistinguishable graphs](#)
[6.6-3 Causal graph of variables and related graphs of values](#)

Definitions

[\(Def. 2.1-1\) Method of rational reconstruction](#)
[\(Def. 2.5-1\) The general scheme of means-end inferences](#)
[\(Def. 2.7-1\) Monotonicity](#)
[\(Def. 2.9-1\) Theory-neutral observation concepts](#)
[\(Def. 3.2-1\) Descriptive versus prescriptive sentences](#)
[\(Def. 3.2-2\) Observational, empirical, theoretical and T-theoretical sentences](#)
[\(Def. 3.3-1\) Logical truth](#)
[\(Def. 3.3-2\) Logical validity](#)
[\(Def. 3.4-1\) Definitional truth](#)
[\(Def. 3.5-1\) Essential sentence types](#)
[\(Def. 3.7-1\) Logical and empirical content](#)
[\(Def. 3.8-1\) Verification, falsification, confirmation and disconfirmation](#)
[\(Def. 3.9-1\) Statistical and subjective probability](#)
[\(Def. 3.9-2\) Conditional probability](#)
[\(Def. 3.9-3\) Principle of narrowest reference class](#)
[\(Def. 3.10-1\) Bilateral reduction sentence](#)
[\(Def. 3.10-2\) Nomological definition of dispositions](#)
[\(Def. 3.12-1\) Relevance in logical inferences](#)
[\(Def. 3.12-2\) Content elements](#)
[\(Def. 3.12-3\) Relevant content](#)
[\(Def. 3.12-4\) Irreducible representation](#)
[\(Def. 4.1-1\) Condition of relevance for strict law hypotheses](#)
[\(Def. 4.1-2\) Statistical relevance for conjunctive antecedents](#)
[\(Def. 4.3-1\) Acceptance interval](#)
[\(Def. 4.3-2\) Significant difference](#)
[\(Def. 4.6-1\) Mean, variance and standard deviation](#)
[\(Def. 5.3-1\) Kinds of theoretical concepts](#)
[\(Def. 5.8-1\) Ramsey sentence](#)
[\(Def. 5.8-2\) Carnap sentence](#)
[\(Def. 5.8-3\) Ramsey-Lewis sentence](#)
[\(Def. 5.9-1\) Classical HD-confirmation](#)
[\(Def. 5.9-2\) Relevant HD-confirmation](#)
[\(Def. 5.9-3\) Incremental and absolute Bayesian confirmation](#)
[\(Def. 5.9-4\) Genuine confirmation \(full and partial\)](#)
[\(Def. 5.9-5\) Goodman's grue predicate](#)
[\(Def. 5.11-1\) Popper-truthlikeness](#)
[\(Def. 5.11-2\) Schurz/Weingartner-truthlikeness](#)

[\(Def. 6.5-1\) Spatiotemporal universality](#)
[\(Def. 6.6-2\) Counterfactual Causality according to Lewis](#)
[\(Def. 6.6-3\) Interventionist causality after Woodward](#)
[\(Def. 6.6-4\) Screening-off between variables](#)
[\(Def. 6.6-5\) Linking-up between variables](#)
[\(Def. 6.6-6\) Causal Markov axiom](#)
[\(Def. 6.6-7\) Faithfulness condition](#)
[\(Def. 6.6-8\) Minimality and productivity](#)
[\(Def. 6.6-9\) Cause-effect relation between values of variables](#)

Key Propositions

[\(Prop. 2.1-1\) The supreme epistemic goal of science](#)
[\(Prop. 2.5-1\) Value neutrality requirement](#)
[\(Prop. 2.7-1\) General pattern of abduction](#)
[\(Prop. 2.7-2\) Inference to the best explanation](#)
[\(Prop. 2.7-3\) Application rule for monotonic inferences](#)
[\(Prop. 2.7-4\) Application rule for non-monotonic inferences](#)
[\(Prop. 2.11-1\) Universal optimality of AW](#)
[\(Prop. 3.1-1\) Translation between logical and set-theoretical language](#)
[\(Prop. 3.1-2\) Linguistic and model-theoretic theory representation](#)
[\(Prop. 3.4-1\) Requirement of empirical non-creativity](#)
[\(Prop. 3.5-1\) Important logical relationships](#)
[\(Prop. 3.9-1\) Basic axioms of probability](#)
[\(Prop. 3.13-1\) Theorems of unconditional probability](#)
[\(Prop. 3.13-2\) Theorems of conditional probability](#)
[\(Prop. 3.13-3\) Binomial formula](#)
[\(Prop. 3.13-4\) Subjective probabilities as coherent fair betting quotients](#)
[\(Prop. 3.13-5\) Statistical principal principle](#)
[\(Prop. 3.13-6\) Weak induction principles for subjective probabilities](#)
[\(Prop. 4.1-1\) Statistical relevance and correlation for qualitative properties](#)
[\(Prop. 4.2-1\) Testing a strict law hypothesis](#)
[\(Prop. 4.3-1\) Testing a statistical law hypothesis](#)
[\(Prop. 4.4-1\) Reichenbach conditions for screening off](#)
[\(Prop. 4.5-1\) Comparative CP law](#)
[\(Prop. 4.5-2\) Exclusive CP assertion](#)
[\(Prop. 4.5-3\) Normic interpretation of eCP laws](#)
[\(Prop. 4.6-1\) Mean value, variance and standard deviation of the distribution of sample means](#)
[\(Prop. 4.7-1\) Principle of Bayes-statistics](#)
[\(Prop. 4.7-2\) Bayesian justification of the likelihood-intuition](#)
[\(Prop. 4.7-3\) Consequences of the indifference principle](#)
[\(Prop. 4.7-4\) Bayesian convergence results](#)
[\(Prop. 4.7-5\) Biased priors](#)
[\(Prop. 5.1-1\) Synthetic nature of laws of correspondence for theoretical concepts](#)
[\(Prop. 5.1-2,3\) Biased indicators](#)
[\(Prop. 5.4-1\) Logical background of the holism of testing](#)
[\(Prop. 5.4-2\) Empirical creativity](#)
[\(Prop. 5.6-1\) Falsification of theory version](#)
[\(Prop. 5.6-2,3\) Successes and failures of theory versions](#)
[\(Prop. 5.6-4\) Confirmation of theory versions](#)

[\(Prop. 5.6-5\) Dynamical truthlikeness](#)
[\(Prop. 5.6-6\) Theory comparison and theory progress](#)
[\(Prop. 5.6-7\) Rational acceptability of theory versions](#)
[\(Prop. 5.6-8\) Untested empirical excess content](#)
[\(Prop. 5.6-9\) Confirmation of theory cores](#)
[\(Prop. 5.6-10\) Preference between theory cores](#)
[\(Prop. 5.6-11\) Rational acceptability of theory cores](#)
[\(Prop. 5.8-1\) Newman's problem](#)
[\(Prop. 5.8-2\) Howson's analysis of the NMA](#)
[\(Prop. 5.9-1\) HD-confirmation and Bayesian confirmation](#)
[\(Prop. 5.9-2\) Impossibility results for rational acceptance](#)
[\(Prop. 5.10-1\) Necessary criteria for transfer of probability increase](#)
[\(Prop. 6.1-1\) DN-explanation of events](#)
[\(Prop. 6.1-2\) DN-explanation of laws](#)
[\(Prop. 6.2-1\) Lawlikeness and causality in DN-explanations](#)
[\(Prop. 6.2-2\) Predictive function](#)
[\(Prop. 6.2-3\) Logical relevance conditions for DN-explanations](#)
[\(Prop. 6.3-1\) Inductive-statistical explanation](#)
[\(Prop. 6.3-2\) Maximal specificity](#)
[\(Prop. 6.4-1\) Normic explanation](#)
[\(Prop. 6.4-2\) Unification condition](#)
[\(Prop. 6.5-1\) Physical necessity](#)
[\(Prop. 6.6-1\) Singular and general causality](#)
[\(Prop. 6.6-2\) Two versions of the causal Markov axiom](#)
[\(Prop. 6.6-3\) Equivalence of productivity and minimality](#)
[\(Prop. 6.6-4\) Causal discovery algorithm](#)
[\(Prop. 6.6-5\) Empirical content of TCN](#)
[\(Prop. 6.6-6\) Causal nets with time-indexed variables](#)

List of Symbols

$\neg$	not
$\wedge$	and
$\vee$	or
$\rightarrow$	material implication (if-then)
$\leftrightarrow$	equivalent (if-and-only-if)
$\leftrightarrow_{\text{def}}$	equivalent by definition
$\forall x$	for all x
$\exists x$	there exists an x
$\exists!x$	there exists exactly one x
$=$	identical
$\neq$	not identical
$=_{\text{def}}$	identical by definition
$\square$	necessary
$\mathbb{N}$	set of natural numbers
$\mathbb{R}$	set of real numbers
$\in$	is an element of

$\{a_1, \dots, a_n\}$	set of individuals
$(a_1, \dots, a_n)$	ordered sequence of individuals
$\{x: Fx\}$	set of all individuals which possess property F
$\subseteq$	proper or improper subset
$\subset$	proper subset
$\cap$	intersection
$\cup$	union
$-$	applied to sets set-theoretical complement
$\times$	Cartesian product between sets
$\cdot$	ordinary product between numbers
$/$	division (example $x \cdot y / z = x$ times y divided through z)
$\square$	empty set
$\sum_i \in_I x_i$ and $\prod_i \in_I x_i$	sum and product (respectively) over all x_i with i in I
D	domain
D^n	n -fold Cartesian product of D
I	interpretation function
$\mathcal{L}$	formal language
$M(\mathcal{L})$	set of models for language $\mathcal{L}$
$\leq$	smaller or equal
$\geq$	greater or equal
$\equiv$	equivalence relation

$f:A \rightarrow B$	function from domain A into set of values B
$\models$	verifies (makes true)
$\Vdash$	follows logically from
$\nVdash$	does not follow logically from
$\dashv \vdash$	is logically equivalent with
p	statistical probability
P	epistemic probability
$C(S)$ and $C_r(S)$	logical content and relevant logical content of a sentence set S
$E(S)$ and $E_r(S)$	empirical content and relevant empirical content of a sentence set S
Ω	possibility space
AL	algebra
$\ A\ $	semantic extension of formula A
$df(x)/dx$	first derivative of f(x)
$\sqrt{\quad}$	square root
$\int$	integral
$\lim_{n \rightarrow \infty}$	limit for n going to infinity
$R(T)$	Ramsey sentence of theory T
$(IN) \text{ DEP } (X,Y)$	variables X and Y are probabilistically (in)dependent

Foreword

I have written *Philosophy of Science: A Unified Approach* for beginners as well as for graduate students and experts. This twofold orientation is reflected in the major structure of the chapters, which start with introductory sections for beginners, followed by sections entitled “complementary and advanced topics,” in which these topics are explained in a condensed way, including presentations of my own accounts. In this way, this book attempts to offer not only an introduction, but also a detailed overview of the various areas of the Philosophy of Science and their interrelations, which will help the reader familiarize him- or herself with the present state of the art in this field.

This book addresses itself both to philosophers and to scientists of all disciplines. By way of bridging Philosophy of Science with neighboring fields such as statistics, hermeneutics, cognitive science, the debate on science and values, and detailed examples of contemporary theories in sciences and humanities, this book tries to convey a deep understanding of the sciences and their methods. This book is *self-contained* and written without presuppositions about the readers. On rare occasions, formulas are used, but I’ve always explained them in a clear and easy-to-follow language.

Although this book introduces the reader to all major areas of philosophy of science, it also has a bias and a general philosophical message. Concerning the bias, the author understands himself as a follower of logically and empirically oriented philosophers of science such as Carnap, Hempel, Reichenbach, and Stegmüller in the German-speaking area. This book’s general message is reflected in its subtitle, *A Unified Approach*. Behind the diversity of phenomena investigated in disciplines as different as elementary particle physics and social sciences, or chemistry and linguistics, one can recognize a deep common ground that unites all of the sciences. This unity, however, is not so much a unity at the level of objects (as it was claimed by the proponents of physical reductionism), but rather a *methodological unity*. [Chapter 2](#) is concentrated on finding the common methodological features of the sciences and their epistemological presuppositions. The remaining chapters will not only work out these common features in all details but will also reveal some important differences between characteristic groups of disciplines, such as the physical sciences versus special sciences (e.g. biology, psychology, or social sciences).

Use in teaching: By restricting the teaching materials to Chapters 1–5 and omitting all sections on advanced topics, one obtains the basis for a semester-long two-hours-per-week course. Covering all of the material contained in this book will fill two such courses. The *core materials* of this book are highlighted in the form of *definitions* and *key propositions*.

Conventions on numberings: The numbering of figures follows the following pattern: “Fig. SecNo.-No.,” where “SecNo.” is the two-digit section number, and “No.” is the number of the figure within the section. Example: “2.1-2” is the second figure of [sec. 2.1](#). In the same way, *definitions* are numbered with “(Def. SecNo.-No.),” *key propositions* with “(Prop. SecNo.-No.),” and *examples* or other *highlighted* passages with “(SecNo.-No.).” Example: “(Def. 4.2-2)” is the second definition of [sec. 4.2](#); “([3.1-4](#))” is the fourth example of [sec. 3.1](#). *Quotations:* Double quotes are used for the purpose of literal quotation and single quotes for stylistic quotation (scare quotes).

At the end of every chapter, the reader finds hints on selected literature and exercises. This book concludes with a comprehensive bibliography and several indices.

This textbook is a revised and significantly extended English elaboration of my German Introduction to Philosophy of Science (“Einführung in die Wissenschaftstheorie”). The introduction has been developed from a series of lectures in the Philosophy of Science that I have given at the Universities of Salzburg, Erfurt, Düsseldorf and, less frequently, at other universities (including California at Irvine, Yale, Vienna, Helsinki, Graz, and Fribourg).

Parts of the text of the German edition have been pretranslated into English by Fergal Treanor. For valuable help concerning matters of content, I am indebted to Ioannis Votsis, Paul Thorn, Hannes Leitgeb, Wolfgang Spohn, Martin Carrier, Franz Huber, Markus Werning, Oliver Scholz, Andreas Hüttemann, Markus Schrenk, Hans Rott, Wolfgang Balzer, Gernot Kleiter, Theo Kuipers, Ilkka Niiniluoto, Erik Olsson, Stephan Hartmann, Jon Williamson, Igor Douven, Ken Gemes, James Ladyman, John Worrall, Stathis Psillos, James Hawthorne, Alan Musgrave, Mathias Frisch, Clark Glymour, Kevin Kelly, Horacio Arló-Costa, and Alvin Goldman. What remains on my side is to wish the readers of this book that it shall bring them much fun and intellectual profit.

Gerhard Schurz
Düsseldorf, March 2013

Foreword to the Second Edition

Since 2013, much has happened in the philosophy of science. Therefore, the Second Edition contains a lot of new content, concerning controversies about the role of values in science (sec. 2.10), results about meta-induction (sec. 2.11), debates concerning the question of scientific realism (sec. 5.8), and the mechanistic philosophy of science (sec. 6.4.1). In compensation, other and more technical parts have been shortened. Moreover, new literature has been added all over the place. The exercises with selected solutions have been extended. The index of subjects has been expanded so that it can be used as a *glossary* : for important terms, the page numbers on which the term is defined are shown in **bold**.

I would like to extend my thanks for valuable content-related help from the first edition to include Karel Lambert, Brian Skyrms, Branden Fitelson, Gila Sher, Johan von Benthem, Lorenzo Magnani, Holger Andreas, Mario Günther, Peter Brössel, Tom Sterkenburg, Leah Henderson, Richard Dawid, Matias Osta Vélez, Jochen Briesen, Gustavo Cevolani, Atocha Aliseda, Petr Spelda, Elke Brendel, Corina Strößner, Alexander Gebharder, Christian Feldbacher-Escamilla, Gottfried Vosgerau, Alexander Christian, Karim Baraghith, Niklas Parwez, Avril Styrman, and Christoph Schamberger.

Gerhard Schurz
Düsseldorf, September 2025

1 Introduction

Where Do We Stand?

DOI: [10.4324/9781003675075-1](https://doi.org/10.4324/9781003675075-1)

This chapter has been made available under a CC BY-SA license.

1.1 Tasks and Aims of Philosophy of Science

Science plays a determining role in the life of today's society. In what sense and to what degree can we trust the results of science? There is a scientific discipline, which examines such questions systematically: the *Philosophy of Science*, which in the German language is called "Wissenschaftstheorie" ("Theory of Science"). Following the German usage of the word, I always understand "science" ("Wissenschaft") in its *broad* meaning – including besides natural sciences, also social sciences, humanities, etc.

The discipline of philosophy (or theory) of science investigates how scientific knowledge works – its goals and its methods, its achievements and its limitations. Like any other scientific discipline, philosophy of science has historically developed out of philosophy. In the present days, it is practiced both by philosophers and by scientists of various disciplines.

One distinguishes between the *general* philosophy of science and philosophies of *particular* sciences. The latter ones are concerned with special kinds of disciplines such as philosophy of physics, biology, psychology, the social sciences, or the humanities. The general philosophy of science discovers those components of knowledge, which are more or less common to all scientific disciplines. Its main questions are:

- i. How is a scientific language constructed?
- ii. What are the rules of correct scientific reasoning or argument?
- iii. What are the distinctive features of a scientific observation or 'datum'?
- iv. What is a scientific law, and what does a scientific theory consist of?
- v. How are law hypotheses and theories empirically tested? When do they count as confirmed, and when as refuted?
- vi. What are the criteria for scientific progress?
- vii. What is achieved by a scientific explanation and what is causality?

This book focuses on these questions of the general philosophy of science, and the following chapters are arranged in corresponding order to the above questions. At the same time, we shall also deal with applications and topics of the philosophies of particular sciences, and examine characteristic *differences* between different branches of science. The most *general* questions of philosophy of science include:

- viii Is there an objectively discoverable reality and an objective truth?
- ix What connection exists between science and value judgments?

Question viii lies on the border between philosophy of science and epistemology (see [sec. 2.8](#)), and question ix on the border with meta-ethics (see [sec. 2.5](#), [2.10](#)).

Beyond its fundamental significance, which consists in providing answers to the above questions, philosophy of science also has a number of significant *applications*, both inside and outside science.

The *internal* applications of philosophy of science consist of the supply of foundational knowledge, which aids the scientist's decisions concerning new or controversial questions. Internal applications also consist of working out *interdisciplinary* common ground and providing *argumentative* competence in critical reflection. Not least, philosophy of science has often paved the way for new scientific disciplines (such as probability theory and statistics, measurement theory, or parts of cognitive science).

Among the *external* applications of philosophy of science to problems in a societal context, two in particular should be emphasized:

- a. The *demarcation* problem is highly significant in society. In this context, it consists of the question of which of our ideas have a claim to the status objective scientific knowledge that should be taught in public educational institutions, as opposed to subjective opinion, political values, ideologies, or religious convictions. This question became politically explosive in the controversy surrounding the creationism movement in the United States. For example, judge W. R. Overton in 1981 (McLean vs. Arkansas) and judge Jones in 2005 (Kitzmiller vs. Dover) based their decisions against the teaching of creationism in school on demarcation criteria between science and religious belief proposed by philosophers of science (cf. [Bird 1998](#), 2–9; [Branch 2011](#), 174). To tell the full truth, there were also philosophers of science, such as Laudan ([1983](#)), who have argued against the possibility of demarcating science from pseudoscience, and they have frequently been quoted by creationists as support for their position (to the embarrassment of other philosophers; cf. [Pennock 2011](#), 176ff).
- b. Equally significant is the critical role of philosophy of science concerning the commercial and political utilization of science, which counteracts the danger of *abusing* science and its results. Politicians, the media, and business leaders enjoy quoting expert opinions, which are presented all too often in a false or one-sided way in order to serve preordained purposes ([sec. 2.10.1](#), [4.3.6](#) – [7](#), [4.4](#)). But also applied scientists are often tempted to put subjective (e.g. financial) interests above objectivity. An example is the discovery of systematic biases in pharmaceutical research toward the financially supporting firms; philosophers of science have promoted scientific journal policies that may help to counteract such biases ([Brown 2008](#)).

1.2 On the Historical Development of Philosophy of Science

Although the term “philosophy of science” was only introduced in the 20th century, the philosophy of science is as old as science itself and had a parallel development (see [Losee 2001](#)). The history of the philosophy of science begins with *Aristotle* (384–322 B.C.), the ancient world's great systematizer of knowledge. Aristotle was much more oriented toward experience than was his teacher Plato. He was, nevertheless, like most philosophers after him, a follower of the *foundationalist* epistemological program. In this epistemological program, it is presumed that real knowledge is only possible if it stands on a foundation of certain and necessary principles, which are arrived at *not* through uncertain experience, but through rational intuition (cf. [Triplett 1990](#), sec. I; [Albert 1985](#), 18–28). Aristotle spoke in this context of “intuitive induction” ([Losee 2001](#), sec. Ia). In the current philosophy of science, by contrast, the *fallibilist* epistemological program has the upper hand, which concedes that our understanding of reality is basically *fallible*, and our scientific knowledge can be more or less well confirmed, but it cannot be guaranteed to be free of error. In a broader sense, of course, the scientific method is foundation-oriented, too, insofar it justifies its knowledge from most evident insights (primarily observation and logic) by inferences that are as reliable procedures possible – in contrast to coherentist epistemological programs, according to which the mutual coherence of beliefs is sufficient for their justification (cf. [Schurz 2024a](#), [sec. 2.1](#) – [2.2](#)).

Considerable advances in science and its philosophy had already been achieved by the late Middle Ages ([Crombie 1959](#)). Under the influence of philosophers such as *Roger Bacon* (1214–1292) or *William of Ockham* (1295–1349), the foundationalist program was gradually replaced by empirically oriented methodology. Of lasting importance is the principle of *Ockham's razor* , according to which one should reject all theoretical assumptions which are not needed to explain the observable. The decisive breakthrough, however, did not take place until the early modern period, the Philosophy of the Enlightenment. Building on the great successes of scientific method in this historical period – thanks to pioneers such as *Galileo Galilei* (1564–1642), *Isaac Newton* (1642–1727), and (later) *Charles Darwin* (1809–1882) – the influential *empiricist* movement was established, beginning with *Francis Bacon* (1561–1626) and *John Locke* (1632–1704) until *David Hume* (1711–1776) and *John Stuart Mill* (1806–1873). In the same period, the foundationalist epistemological program developed into the *rationalist* movement, prominently represented by *René Descartes* (1596–1650), *Gottfried W. Leibniz* (1646–1716), and *Immanuel Kant* (1724–1804).

The essential difference between the empiricist and rationalist philosophies can be characterized as follows. For the empiricists, sentences which may be justified *a priori* – that is, by pure reason and with rational certainty – are limited to so-called analytical sentences, that is, sentences whose truth is based on logic and conceptual convention. Such sentences possess no realistic content; they say nothing about the real world – only *synthetic* sentences may do this, but they can only be justified *a posteriori*, that is, based on experience. For the rationalists, on the other hand, *a priori* justifiable sentences with realistic content do exist, namely the so-called synthetic *a priori* sentences. However, both Descartes' and Kant's attempts to justify synthetic *a priori* principles of the sciences were unsuccessful and have been refuted by the further development of science. The development of the modern philosophy of science is closely bound with the insight that the foundationalist epistemological program was not tenable – at least not in its 'strong' version as explained above.

For a long time, classical systems of philosophical empiricism too had reductionist and foundationalist tendencies. The skeptical consequences of empiricism, which called for epistemic humility, were first recognized and radically formulated by Hume. Hume showed that the two central tenets of scientific method, the *causality* and the *induction principle* , are neither logically nor empirically justifiable. The causality principle says that all regular successions between events are produced by cause–effect relations; but Hume objected that all what one can observe are the successions of the events, while the idea that the earlier event or 'cause' produces the later event as its "effect" is metaphysical fiction (see [sec. 6.6.2](#)). The induction principle asserts that regular connections between events that have been observed in the past will also hold in the future; but Hume argued that this inference can neither be justified by logic nor by experience (see [sec. 2.11.1](#)). The problems involved in the justification of the principles of induction and causality have occupied philosophers up to the present day.

The apparent *insufficient justifiability* of empiricism put wind in the sails of subsequent rationalist movements, in which the foundationalist epistemological program, however, was gradually given up or replaced by pragmatized or historicized 'rationalisms' (see [sec. 1.3.2](#) –1.3.3). In the 20th century, *post-empiricist* and *post-rationalist* approaches have moved considerably closer to each other, and in the context of this tension, present-day philosophy of science has developed.

1.3 Philosophical Positions in Contemporary Philosophy of Science

1.3.1 Logical Empiricism

One of the most important origins of modern Philosophy of Science and Analytic Philosophy is logical empiricism, which was developed in particular by the Vienna Circle, but also by others, for example by members of the Berlin circle around Hans Reichenbach (1891–1953). The *Vienna Circle* was a group of philosophers and scientists in Vienna, whose core consisted of Moritz Schlick (1882–1936), Otto Neurath (1882–1945), and Rudolf Carnap (1891–1970). In the tradition of Ernst Mach (1838–1916), this group attempted to reestablish empiricism and scientific philosophy in general (cf. [Stadler 2001](#)). What was new about their situation was the development of *modern logic*, which had not been founded until the end of the 19th century. Modern logic made it possible to represent any language-based epistemological system with mathematical precision, and so the Vienna Circle hoped they had found at last the methodological equipment for a philosophy which could progress in a scientific manner (cf. [Schlick 1930/1931](#), 155f). In the middle of the 1930s, the Vienna Circle, with political leanings from liberal to socialist, had to disband under pressure from the National Socialists. The majority of its members emigrated, some via other countries, to the United States, where the logical-empirical movement, uniting with related Anglo-American directions, developed into Analytic Philosophy and gained a foothold in Europe in the postwar period.

Even today, many textbooks present logical empiricism as if it had represented a single dogmatic position; in reality, however, its proponents were highly capable of learning and continuously revised their positions. What today's philosophy of science can learn from logical empiricism is not so much any particular thesis but rather the high standards of conceptual and argumentative precision. In the phase up to 1935, the philosophy of science of logical empiricism was restricted by *positivism* and *reductionism*. In the later phase, the logical empiricists gradually rejected their restricted positions and replaced them with theses, to which the designations “empiricism” or “positivism” in the usual sense *no longer apply*. In one regard, logical empiricism has parted company with its classical predecessors early on: with the refutation of the classical empiricist thesis that at least ‘direct’ observation reports are infallible. In the so-called protocol sentence debate, Neurath and Carnap – but not Schlick – became convinced that even observation sentences such as “there is a table there” are in principle fallible. Empiricism, according to this new view, just meant that observation sentences have an epistemically favored role within the whole system of discovery (see [Carnap 1932/1933](#); [Neurath 1934](#), 113).

Empirical Reductionism, by contrast, was retained for quite some time. Classical empiricism advocated the following reductionist thesis: all ‘serious’ scientific concepts must be translatable via a chain of definitions into observational concepts. This program of reduction was adopted by logical empiricism in its early phase ([Carnap 1928a](#)). Later, however, the view prevailed that *theoretical concepts* such as “power,” “electrical field,” or “human character” *cannot* be defined by observational concepts, but rather go far beyond the immediately observable ([Hempel 1951](#); [Carnap 1956a](#)). In the so-called *standard* philosophy of science of the 1960s, the classical empiricist epistemological model had already broken down. What remained was a *minimal empiricism*, which consisted of the requirement that scientific theories must have empirical consequences, by means of which they can be tested.

A similarly modest empiricism is defended by contemporary empiricists such as van Fraassen (1980, [2002](#)). He does not only reject infallibilism and reductionism but also the *linguistic* orientation of logical empiricism. Instead, he defends a semantic conception of theories, in which theories are regarded not as sets of sentences, but as sets of conceptual models, similar as in structuralist philosophy of science (cf. [sec. 3.1.2](#)). Characteristic of van Fraassen's empiricism (1980, sec. 1) are two ingredients: first, a defense of the concept of theory-independent observation against Kuhn-followers (see [sec. 1.3.3](#)). And second, his rejection of scientific realism in favor of an instrumentalist position, according to which the theoretical (i.e. non-observable) part of a theory can only be reasonably said to be more or less *empirically adequate*, but not true or false. In this latter respect, van Fraassen is more restrictive than Carnap (e.g. 1928a) who always argued for ontological

pluralism , according to which both realism and instrumentalism are possible but epistemically undecidable practical attitudes toward scientific theories.

Two boundaries , which early logical empiricism had tried to maintain, became permeable in contemporary ‘minimal’ empiricism. *First* , the boundary between *science* and *metaphysics* became porous: it turned out that principles, which are of no empirical consequence *in isolation* , may give rise to new empirical consequences when joined with other theoretical sentences (see [sec. 5.4.2](#) ; [Hempel 1951](#) , 115; [Stegmüller 1969](#) ff, vol. II.1, 293–295). In this way, the *post-positivist* philosophy of science gained new access to the discussion of metaphysical principles such as realism and causality. Conversely, “inductive” branches of metaphysics emerged that sought to justify their principles not by pure thought but through empirical science (cf. [Hüttemann/Schurz 2025](#)). *Second* , the boundary between analytic and synthetic sentences became porous when it was recognized that the meaning of theoretical terms is not determined by isolated definitions, but by the respective background theory as a whole ([Carnap 1956a](#) , [1963](#)). In parallel fashion, Quine ([1951](#)) argued that in *natural* , empirically observed languages, no clear criterion exists by which one could undertake a demarcation between analytic and synthetic sentences.

As a consequence of this, analytic philosophy’s *linguistic* orientation (e.g. [Carnap 1950a](#)) was gradually given up. Linguistically oriented philosophers try to answer all philosophical questions by means of semantical concept analysis. But if a clear divide between (analytic) questions of meaning and (synthetic) questions of fact is missing, this exclusively linguistic orientation is no longer tenable. For this reason, Carnap ([1956b](#) , suppl. E) started to work on a system of pragmatics. Later on, Quine ([1974](#)) developed a *naturalistic* view, according to which science and philosophy form a continuum, the latter’s theses also having to be corroborated by experience.

1.3.2 Critical Rationalism

Also associated with the Vienna Circle was *Karl Popper* (1902–1994), the founder of so-called critical rationalism. From the beginning, Popper’s theory of science displayed the very non-reductionist tendencies that logical empiricism had accepted only after years in transition. According to Popper, scientific theories may go any distance beyond our experience, as long as they are empirically testable. Popper’s tests consist of *falsification attempts* , which should be as severe as possible. Popper supported his emphasis on falsification (instead of verification) by citing the *asymmetry* of *verification* and *falsification* of strictly universal hypotheses such as “all metals conduct electricity.” Such hypotheses cannot be verified by any finite number of observations, because they speak about a potentially infinite domain of instances – in our example, all metals in the past and in the future. However, these hypotheses can be falsified by a single observed exception ([Popper 1935](#) /2002, sec. I–IV).

For Popper, falsifiability was the decisive criterion of *demarcation* between science and speculation: scientific theories are not verifiable, but they must be falsifiable. Once falsified, they are *rejected* as false; but if they withstand manyfold falsification attempts, they are viewed as *corroborated* .

Popper’s falsificationism was also subjected to diverse criticism in the discussion which followed. Lakatos ([1970](#)) pointed out that scientific theories are almost never rejected on the basis of one single counterexample, but may be immunized against conflicts with recalcitrant data by means of modifications that are usually more or less *ad hoc* (see [sec. 5.6.1](#)). Popper’s *anti-inductionism* – that is, his thesis that science could progress without making any use of induction – also came in for harsh criticism (see [sec. 2.7.2](#)).

Popper’s critical rationalism is far removed from classical rationalism. Nowhere in Popper’s philosophy is it claimed that knowledge of reality can be grounded in non-empirical a priori intuition; on the contrary, Popper always rejected this. Thus, Popper’s philosophy is *post-rationalist* just as that of the late logical empiricists is *post-positivist* . Popper stressed more strongly than late logical

empiricism, however, that observation sentences are not only fallible, but also *theoretically biased* , so that the borderline between observational and theoretical terms cannot be sharply drawn (1935/2002, sec. V.30 addendum; new app. X). With this argumentation, critical rationalism opened its doors to the attacks from historical relativism, to which we turn now.

1.3.3 The Historical Account and Relativism

At the end of the 1950s, the historian of science Thomas Kuhn ([1962](#)) rattled the scientific world by presenting a fundamental attack on the standard philosophy of science. Factual science, said Kuhn, behaves differently from the normative ideals ‘prescribed’ by the philosophers of science. The alternative scientific model developed by Kuhn takes more of a historical-sociological than a logical-cognitive approach. According to Kuhn, scientific development takes place on the basis of so-called paradigms, such as the paradigm of classical (Newtonian) physics or (Darwin’s) theory of evolution. A Kuhnian paradigm contains at least the following three components: (i) very general theoretical principles or models, (ii) typical examples of successful applications, and (iii) methodological-normative assumptions (cf. [Kuhn 1977](#) ; [Hoyningen-Huene 1993](#) ; [Schurz 1998a](#) , [2014c](#)). The sociological correlate of a paradigm is a *scientific community* , a community of experts in the given field, who stick to a paradigm and work on its further development. According to Kuhn, a paradigm determines not only the fundamental principles, but also the major open problems and the *interpretation* of observed data – no, it even determines the observed data *themselves* , because, according to this view, *all observation is theory-laden* , and hence theoretically biased. For Kuhn, there is no observation which is independent of background theory or paradigm. Kuhn takes this strong thesis from Hanson ([1958](#)).

According to Kuhn, scientific development takes place in two successive phases, a *normal* and a *revolutionary* phase. Common acceptance of a paradigm makes the continuous growth of knowledge possible in the normal phase of a scientific discipline. When recalcitrant data, so-called anomalies, resist explanation by the given paradigm, the conflict is settled by more-or-less *ad hoc* modifications of current versions of theories within the accepted paradigm. Should such anomalies become more frequent, however, younger scholars will begin to look for a *new* paradigm. As soon as one is found, scientific development enters a revolutionary phase for a certain time, in which two paradigms struggle for supremacy. Kuhn’s examples in (1962) include the transition from Ptolemaic to Copernican astronomy and from Newtonian to Einsteinian physics. Kuhn claims that there are no common standards of rationality during a paradigm change and all empirical data are seen by the new paradigm in a new way. For this reason, the two competing paradigms are, according to Kuhn’s ‘notorious’ *incommensurability thesis* , rationally incomparable. So the struggle for supremacy does not take place in the form of a comparison of cognitive achievement, but rather in the form of a *power struggle* , in which supporters of the old paradigm eventually die out, as a result of which the new paradigm prevails, and a new normal scientific phase is heralded in.

The standard philosophy of science was unable to effectively counter the argument of the theory-dependence of scientific observations and its relativistic consequences. For all questions about the cognitive nature of observation had been banned by the narrowed perspective of the *logic* of science. For Popper, the question of what should be called an observation sentence was a question of pure convention (1935/2002, sec. V.30). The logical empiricists saw it in much the same way. It will be shown in [sec. 2.8.2](#) that a different picture results as soon as we take on an *empirical-cognitive* perspective. After all, if there were no difference between what experience *shows* , and what theory leads us to *expect* , then empirical science would be an exercise in *permanent self-deception* . Any theory of the empirical sciences which cannot explain this central difference must be extremely unsatisfying. Moreover, the loss of a viable criterion of empiricism would also undermine any credible distinction between real and fake facts. After right-wing movements began to criticize scientifically established facts, for example, concerning the climate crisis, as “leftist ideology,”

philosophers with left-leaning views tend to embrace once again the experience-based objectivity that was criticized by them some decades ago as an expression of “positivist ideology.”

Kuhn’s alternative model of science was the subject of much controversial discussion in the decades that followed. As well as an opposition to Kuhn, both a moderate and a more radical branch of Kuhn-followers was formed (cf. [Hoyningen-Huene 1993](#) , 212f). The more radical group gave relativism a strong impetus – most prominently in Feyerabend’s “Anarchist Epistemology” (1975), who with his maxim “anything goes” argued against methodological rules that are guiding or should guide science. The moderate Kuhnians, by contrast, aimed to separate Kuhn’s insights in the philosophy of science from the exaggerations of his “radical interpretation.” Kuhn himself later softened his incommensurability thesis and granted limited recognition to the possibility of rational comparisons between paradigms. ¹A moderate Kuhn-follower, and a Popper-follower at the same time, is Imre Lakatos and his conception of scientific research programs. Another moderate Kuhn-follower is Larry Laudan ([1981](#)), who argued that the empirical success of a scientific theory is not necessarily a reliable indicator of its truth. Such moderately Kuhnian developments have brought important insights to light for the philosophy of science (see [sec. 5.6](#) and 5.8). A further controversy centered around Kuhn’s thesis that in its normal scientific phases a scientific discipline is always dominated by a single paradigm. Scientists and philosophers of science have pointed out that, contrary to Kuhn’s thesis, disciplines outside of physics often exhibit a persistent coexistence of rival paradigms (cf. [Schurz/Weingartner 1998](#) ; [Kornmesser/Schurz 2014](#)).

1.3.4 Metaphysical and Scientific Realism

In reaction to the Kuhnian challenges of the logical-empirical model of science, many philosophers have suggested returning to the metaphysical approaches of realism and necessity, from which the philosophy of science in the early 20th century had tried to free itself. For example, it is a consequence of post-positivist philosophy of science that the term “mass” actually has different *meanings* in Newtonian and Einsteinian physics. This theory-dependence was also one of the main pillars of Kuhn’s incommensurability thesis. The early Putnam (e.g. 1975b, 219ff, 229–235, 240f), together with Kripke ([1972](#) , 55–59), developed the following argumentation against this consequence: scientific progress in the sense of a movement toward objective truth is only possible if the *reference* of scientific terms, that is, their relationship to reality, *is rigidly fixed* with metaphysical necessity, in the sense that this reference stays the same in all possible theories and across all possible worlds. Kripke and Putnam assumed the reference fixation to take place by acts of baptizing, together with a rigid sameness-relation between natural kinds, with the result that statements such as “water is H₂O” become true by “metaphysical necessity” (see also [Shoemaker 1980](#)). Fine ([1994](#)) also regards such identity relations as metaphysical necessities constituted by essential properties.

We agree with Kripke and Putnam that the meaning of natural kind terms such as “water” cannot be learned by an empirical definition (as in [Carnap’s 1956b](#) account), but by acts of ostension (“this is water”) grasping the right sameness-in-kind relation between water exemplars. But as a number of arguments have shown, scientific knowledge cannot determine sameness-in-kind relations and the reference of (theoretical) concepts with metaphysical necessity. ²Of course, classification systems for scientific (e.g. chemical or biological) kinds can be fixed *analytically* – that is, through definitional conventions (cf. [Sidelle 1989](#)). However, which classification system accurately reflects reality depends on the theoretical background knowledge ³which undergoes historical change ([Hendry 2008](#)). The idea that a necessary relationship between language and experience can be established independently of experience is in the end a regression toward a foundationalist epistemic model. In fact, the later Putnam ([1977](#) , [1982](#)) relinquished his earlier metaphysical realism by arguing as follows: if neither our empirical nor our theoretical knowledge can determine the reference of our concepts in a unique way, then *nothing* can do this – the consequence is an

unavoidable “indeterminacy of reference.” The same indeterminacy of reference has been worked out by Quine, but while Quine remained a naturalistic realist, the late Putnam adopted a more-or-less idealistic position that he called “internal realism.”

An account of necessity that is metaphysical in a weaker sense than the above ones is that of Armstrong ([1983](#)). According to him, laws of nature supervene on an assumed necessitation relation between “universals” (such as “water” and “H₂O”), which is, however, *contingent* – that is, these necessities may be different in different possible worlds ([sec. 6.5.6](#)). A similar middle ground between metaphysics and empiricism is Lewis ([1973](#) , [1986a](#)), who developed a possible worlds semantics for counterfactual and causal conditionals, but at the same time defends an elaborated Humean account to laws of nature, according to which laws supervene on the ‘mosaic’ of empirical facts ([sec. 6.6.5](#) ; cf. [Schrenk 2017](#) , [sec. 4.3](#) , [5.3](#)).

Like metaphysical realists, proponents of scientific realism also advocate for the convergence of science toward truth – but not on the basis of metaphysical postulates, but based on the empirical success of science. To support this position, scientific realists require robust arguments that allow them to infer the approximate truth of a theory from its empirical success. The most well-known argument of this kind is Putnam’s *no miracles* argument (1975a, 73). According to this argument, the empirical success of a theory would be a sheer miracle unless one assumes that the theory is at least approximately true. The no miracles argument has been subject to extensive criticism, notably by Laudan ([1981](#)), who, in his *pessimistic meta-induction* , argued that virtually all empirically successful theories of the past have eventually been shown to be false, whence it is inductively highly probable that our current best theories will suffer the same fate. Other philosophers of science, such as Boyd ([1984](#)) and Psillos ([1999](#)), have defended scientific realism against such anti-realist arguments, leading to a scientific realism debate that continues to this day (see [sec. 5.8.3](#) – [5.8.8](#)).

1.3.5 Naturalism and the Cognitive Turn

Modern naturalism goes back primarily to Quine ([1951](#) , [1960](#)), who rejected the analytic-synthetic distinction, thereby dissolving the clear boundary between scientific philosophy and empirical science. For Quine, naturalized epistemology becomes an empirical discipline, which is “scientifically investigating man’s acquisition of science” ([Quine 1974](#) , 3). The naturalistic program was emphatically taken up by a number of young epistemologists and philosophers of science (e.g. [Kitcher 1992](#) ; [Papineau 1993](#) ; [Kornblith 1994](#)).

Philosophical naturalism found expression in several philosophical programs. A first concretization of naturalism is *evolutionary epistemology* (e.g. [Campbell 1984](#)): many aspects of humans’ cognitive capacities and achievements can be explained by facts of evolution, in particular regarding perceptual knowledge. It has to be critically remarked, however, that naturalistic approaches themselves make certain epistemological assumptions which are *in need of justification* , for example concerning realism and induction, and so they cannot entirely replace classical epistemology. There is, however, an epistemological school of thought that indeed amounts to a *renunciation* of justification in this regard, namely epistemological externalism. According to its most well-known version ([Goldman 1986](#)), a belief counts as justified if it is produced by an “objectively reliable” method, without requiring anyone to be able to justify the reliability of the method in question. A justification of inductive prediction methods is not sought; instead, it is simply asserted that a predictive method is justified if it is reliable, that is, if its predictions will come out as true. Otherwise, it is considered unreliable. The externalist treatment of realism works in a similar way. However, since all nonscientific ‘methods’ – ranging from creationist doctrines of salvation to clairvoyance – also claim reliability for themselves, externalist epistemology fails to demarcate science from nonscience and thereby condemns itself to epistemic uselessness (see [Schurz 2024a](#) , [sec. 2.4](#)).

A second concretization of naturalism is the so-called cognitive turn. This turn resulted from the fact that meanwhile the cognitive sciences had attained the status of an interdisciplinary subject area in their own right, in which the basis of humans' cognitive and scientific capabilities is examined not only logically in philosophy but also empirically in psychology or linguistics and computationally in Artificial Intelligence research (cf. [Thagard 1996](#) ; [Anderson 2000](#)). As a result, several philosophers have started efforts to bring the approaches of the cognitive sciences into philosophy of science and vice versa (e.g. [Goldman 1986](#) ; [Bechtel 1988](#) ; [Bishop/Trout 2005](#)). The recent movement of “experimental philosophy” has even started to examine philosophical intuitions by empirical investigations instead of the traditional “armchair method,” with interesting results concerning cultural differences ([Machery et al. 2004](#) ; [Knobe/Shau 2008](#) ; [Horvath/Koch 2020](#)).

The third specification of naturalism is, finally, naturalistic philosophy of science. Like Quine, it assumes a continuity between the methods of the empirical sciences and those of the philosophy of science. Questions of value are not excluded from science, but they are understood in an instrumental way – as questions of the optimal choice of means for given ends ([Laudan 1987](#)) – which closely aligns with the value-neutrality approach discussed in [sec. 2.5](#) . By contrast, the influence of externalism must be viewed critically: the philosophical aim of justifying fundamental scientific methods in contrast to pseudoscience is largely abandoned, and the reliability of scientific methods and results is simply presupposed ([Godfrey-Smith 2003](#) , 151). This runs counter to the aim of theoretical clarification of science and the exposure of the ideological misuse of scientific authority, as outlined in [sec. 1.1](#) .

1.3.6 Structuralist and Semantic View of Theories

In search of a better defense against the Kuhnian offensive, Stegmüller ([1976](#) ; [1977](#) , vol. II., sec. 6; 1979b) suggested replacing the “statement view” of the standard philosophy of science with a so-called non-statement view. Stegmüller's suggestion is based on the idea, traceable to Suppes ([1957](#) , 246ff) and Sneed ([1971](#)), that it is better to reconstruct scientific theories as systems of set-theoretical *models* than as systems of statements in a logical language. Such systems of models are also called *structures* : hence, the name “structuralist” philosophy of science. In 1976, [Stegmüller \(1976](#) , 105) envisions this step as a radical upheaval. Later, he is more modest and argues for the transition from logical to set-theoretical formalization methods for “pragmatic and psychological reasons” concerning the simplicity of formal reconstructions (1979b, 65). For set-theoretical models are also characterized by statements, namely by statements in the language of set theory. Therefore, there exist straightforward possibilities of *translation* between the theory representations in the statement view and the non-statement view (see [sec. 3.1.2](#) ; [Schurz 2014b](#)).

While the statement-versus-non-statement debate turned out to be a bit of a red herring, the structuralist philosophy of science can claim significant achievements in the *detailed reconstruction* of individual scientific theories and theory nets (cf. [Balzer et al. 1987](#) ; [Balzer/Moulines 1996](#)). Moreover, the structuralist philosophy of science has received support from Anglo-Saxon philosophers such as van Fraassen (1980, sec. 2; 1989, sec. 9) and Giere ([1988](#) , [2004](#)), who called this position the “semantic view of theories” and developed it in a less formal way (cf. [Suppe 2000](#)). The semantic view of theories essentially consists in the claim that theories should be understood as models – a view later taken up by the “new mechanists” in the philosophy of science (e.g. [Glennan 2017](#) ; [Glennan et al. 2022](#) ; see [sec. 6.4.1](#)). However, the concept of a model remains vague: virtually anything can serve as a ‘model’ – from images to mathematical equations – as long as it serves to represent some aspect of reality (see [Glennan 2017](#) , 60; for criticism, see Chakravartty 2001; [Schurz 2014b](#)).

1.3.7 Pragmatic Philosophy of Science

In the 1980s, several authors spoke of a *pragmatic turn* in philosophy of science ([Stegmüller 1969](#) ff, vol. I, sec. XI; [Böhler et al. 1986](#) ; [Bernstein 2001](#)). Unfortunately, however, pragmatic philosophy and the use of the term “pragmatics,” in general, fall into two *opposing* camps. While one camp (e.g. Ch. S. [Peirce 1931](#) ff, van Fraassen 1980; [Rescher 1998](#) , [2001](#) ; [Schurz 1998](#) b) understands pragmatics in the *epistemically internal* sense, the other camp (e.g. [James 1907](#) ; [Rorty 1982](#) ; [Stich 1990](#) ; partly [Putnam 1995](#) ; Bernstein 2010) sees pragmatics in an *epistemically external* sense.

For the epistemically external camp, the pragmatic components of discovery are essential, but have nothing to do with its truth. By contrast, for the epistemically internal camp, it is these very pragmatic components, which are supposed to establish an objective conception of truth in a noncircular sense. The epistemically external camp makes knowledge (or justified belief) dependent on *any* kind of interests, be it predictive accuracy, political control, or emotional well-being, and sees therein a compelling argument for the unavoidable subjectivity of knowledge. In this sense, Stich ([1990](#) , 131ff) has argued that there are so many different conceptions of truth as there are different cognition-guiding interests. By contrast, for the internal camp, the reference to epistemically *internal* purposes is the very thing which makes possible a noncircular foundation of knowledge. For example, the reference to the cognitive purpose of maximizing true predictions enables a justification of induction that escapes Hume’s skeptical challenge ([sec. 2.11.2](#)). In conclusion, the external and the internal camp *differ so greatly* in their understanding of pragmatics, that they should not be subsumed under a single term.

1.3.8 Hermeneutics and Critical Theory

The *Philosophy of the Humanities* , which originated in the 19th century, goes back to Friedrich Schleiermacher (1768–1834), Wilhelm Dilthey (1833–1911), and Wilhelm Windelband (1848–1915). It started as a movement for the methodological demarcation of the humanities against the expanding natural sciences. Central to this philosophy is the use of *hermeneutics* , the study of interpersonal *understanding* (*Verstehen*) , for the purpose of such a methodological demarcation. The *methodological dualism* of the philosophers of the humanities has been brought to a head with the following thesis: in the natural sciences, we *explain* , but in the humanities, we *understand* . While scientific explanations are based on general law hypotheses, it was claimed that understanding in the humanities needs no recourse to such law hypotheses, because the human mind is not subject to strict law mechanisms, and understanding has direct access to the human mind.

Hermeneutics is part of a tradition from well before Schleiermacher, in which there was no mention of any rigid separation of the humanities from logic and natural sciences (see [Scholz 2001](#) , part I.B). The methodological dualism originated in the 19th century and had a strong influence on methodology debates of the 20th century (cf. [Von Wright 1971](#) ; [Manninen/Tuomela 1976](#) ; [Apel 1984](#)). The contemporary controversy in the humanities and social sciences about *quantitative* versus *qualitative* methods is partially a revivification of the old methodological dispute (cf. [Lamnek 1988](#) ; [Denzin/Lincoln 2000](#)). In response, the analytic philosophy of science has produced a number of approaches, in which the unity of hermeneutics and empirical methods is worked out (see [sec. 6.4.2](#) ; cf. [Hirsch 1973](#) ; [Follesdal 1979](#) ; [Taylor 1980](#) ; [Bühler 2003](#)).

Critical theory was developed by the philosophers of the so-called Frankfurt school, mainly Horkheimer, Adorno, Marcuse, and Habermas. It combines the methodologically dualist version of hermeneutics with the doctrine of the unavoidable interest-dependence of discovery and knowledge (see [Habermas 1966](#) ; [Apel 1984](#)). Critical theory also influenced the philosophy of science debate on the value-freedom versus value-ladenness of scientific research. We turn to this question in [sec. 2.5](#) and 2.10, where a modernized version of scientific value neutrality is developed – one that is compatible with value commitment in science and capable of withstanding relevant objections.

1.3.9 Radical Constructivism, Postmodernism, and the Science Wars

Radical constructivism was developed in the 20th century by biologists such as Maturana and Varela ([1992](#)) and psychologists such as von Glasersfeld ([1995](#)); philosophically, however, it is much older. The central argument of constructivism infers that, because our perceptions and descriptions of reality are constructed, there is no subject-independent reality – instead, *reality itself* is said to be a human construction. Even the early Kuhn embraced a constructivist position, claiming that “the world itself” has changed following a scientific paradigm shift ([Kuhn 1962](#) , chapter 10; see [Godfrey-Smith 2003](#) , 96).

In [sec. 2.8.2](#) , we will show that the argumentation of radical constructivism is based on a *fallacy* . Regardless of this, radical constructivism, in combination with the hermeneutic methodological dualism, has exerted its influence in the cultural and social sciences. There, it shaped French poststructuralism and the philosophy of postmodernism ([Lyotard 1984](#)), which reject scientific reality and its impact on human culture, instead promoting the idea that reality is entirely socially constructed. In the United States, these currents gave rise to the field of “Science Studies” and encouraged a left-leaning politicized understanding of science, as exemplified by feminist philosophy of science ([Harding 1986](#) ; cf. [Godfrey-Smith 2003](#) , sec. 9.1–9.4). Most pointedly, the social constructivist position can be found in the “science studies” of Latour and Woolgar ([1986](#)) and Knorr-Cetina ([1981](#)), where scientific laboratories (e.g. in nuclear physics or biochemistry) are studied by ethnographical methods of studying “foreign cultures” without engaging with the actual epistemic content being investigated, in order to learn something about the “social construction of scientific reality.”

The disregard of methodological standards and the misuse of scientific terminology met with strong opposition from the scientific community and led, in 1996, to a “counterattack” in the form of the *Sokal Affair* . In order to expose the lack of qualifications behind the wordplay of postmodern authors, Sokal ([1996](#)) published a hoax article titled “Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity,” which consisted of a jumble of quotes from postmodern thinkers. Despite its nonsensical content, the article was accepted and published by the postmodern-oriented American journal *Social Text* . When Sokal later revealed in another journal that his article had been a parody of postmodern discourse, the scandal was complete (cf. [Godfrey-Smith 2003](#) , 146f). The subsequently published book by Sokal and Bricmont ([1998](#)), entitled *Fashionable Nonsense: How the Postmodernists Abuse Science* , caused a major stir and became a catalyst for the so-called Science Wars ([Ross 1996](#)).

1.4 Introductory Literature and Exercises

1.4.1 Introductory Literature

A historically oriented introduction into the philosophy of science of the 20th and 21st centuries is Godfrey-Smith ([2003](#)). Ladyman’s introduction (2002) is focused on natural science and Rosenberg ([2008](#)) introduces to the philosophy of social science. Okasha ([2016](#)) gives a systematic but very brief introduction. An excellent introduction to the history of the philosophy of science is offered by Losee ([2001](#)). Stadler ([2001](#)) presents an extensive study of the Vienna Circle and its logical empiricism. For Popper’s philosophy of science, see [Keuth \(1998\)](#) , ed.). Hoyningen-Huene ([1993](#)) offers an introduction into the philosophy of Kuhn. [Hüttemann \(2021\)](#) offers an introduction into the metaphysics of science; for critical discussion of metaphysical accounts, see Ladyman and Ross ([2007](#)). On evolutionary epistemology, see Campbell ([1984](#)). The history and current situation of hermeneutics are dealt with by Bruns ([1992](#)).

1.4.2 Questions and Exercises

Sec. 1.1–1.2

1. What are the seven main questions in the general philosophy of science?
2. What is the major difference between the empiricist and rationalist ways of thinking in epistemology?

Sec. 1.3

1. What is understood by empirical reductionism?
2. What criterion of demarcation is suggested by critical rationalism?
3. What is a paradigm for Kuhn? Discuss Kuhn's incommensurability thesis.
4. How does scientific instrumentalism differ from scientific realism?
5. Discuss the methodological dualism of the philosophers of the humanities from a contemporary viewpoint.

Notes

1. Kuhn ([2000](#) , 33–57); cf. Okasha ([2016](#) , 87); [Seselja/Straßer \(2013](#) , 14f). Empirical data show that history of science proceeds much more cumulative than Kuhn suggests (see [Krauss 2025](#)). ↩
2. See [sec. 5.8.2](#) . Cf. Mellor ([1980](#) , 118f); Bird ([1998](#) , 108–120); Ladyman/Ross ([2007](#) , 290ff). ↩
3. In philosophy of science, “background knowledge” is not understood in the strict epistemological sense: it contains scientific beliefs that are rationally accepted but not necessarily true. ↩

2 The Question of Unity

In Search for Common Foundations of the Sciences

DOI: [10.4324/9781003675075-2](https://doi.org/10.4324/9781003675075-2)

This chapter has been made available under a CC BY-SA license.

2.1 Normative or Descriptive? The Method of Rational Reconstruction

Prima facie, we can identify two opposing views of the goals and methods of the philosophy of science:

1. According to the *normative* view, it is the goal of philosophy of science to say what science *should* be, and how it should be pursued. For this purpose, philosophy of science must determine in what scientific rationality consists, and on the basis of which criteria a scientific hypothesis may be justified.
2. According to the *descriptive* view, on the other hand, it is the task of philosophy of science to say what science de facto *is*, and how it proceeds. For this purpose, philosophy of science must describe and explain the existing sciences in terms of their historical development and their current structure.

The normative view is historically older, and currently still enjoys a wide following. Both logical empiricists and critical rationalists held this view. It was not until Kuhn ([1962](#)) set off the historical turn in philosophy of science that the descriptive counterpoint became popular. It can be found in the work of historically oriented philosophers of science (cf. [McMullin 1970](#) ; [Laudan 1977](#) , 162), of the structuralists (e.g. [Stegmüller 1976](#) , 267ff; [Balzer et al. 1987](#) , xvff), and of the naturalists (cf. [Giere 1999](#) , 160–163; [Ladyman 2002](#) , 4). So which position is right?

The position of the normativists can be traced back to Hans Reichenbach's ([1938](#) , 6f) distinction between the *context of discovery* , that is, the factual *genesis* of scientific knowledge, and its *context of justification* . The plausible prima facie argumentation of the normativists may be broken down as follows (e.g. [Popper 1935](#) /2002, [sec. 1.2](#)):

(2.1-1) *Normativist argumentation* :

Premise 1 : To assess the epistemic value of a hypothesis, *only* the question of rational justifiability of its truth claim is relevant.

Premise 2 : Questions about the hypothesis's historical development, the causes for its discovery, etc., are irrelevant to this purpose.

Conclusion : Therefore, philosophy of science should only examine science's rational justification and not its context of genesis – the latter being the remit of the history, sociology, or psychology of science, but having nothing to do with its philosophy.

Since Kuhn ([1962](#)), the argumentation of the normativists has been criticized in a very *sweeping* way. Here, we will attempt to clarify what is correct and what is flawed in this line of argument. The argument is certainly logically correct – assuming we take it to be analytically true that the philosophy of science is concerned with the assessment of the epistemic value of hypotheses. That is to say, *if* the premises of the argument were really true, then the conclusion would also be true. There is also *no error* in premise 1 – this thesis also seems to be analytically true, to the extent that science is concerned with the discovery of truth, and nothing else. But premise 2 is already infected by the 'germ' of the normativist fallacy. Because the

historical genesis of a scientific hypothesis includes not only so-called external factors which are epistemically irrelevant, but also internal factors which are epistemically relevant: for example, the experimental methods, which scientists de facto use, and the rational hypothesis evaluation, which they de facto apply.

The organic chemist Kekulé, for example, reported that his groundbreaking hypothesis of ring-shaped benzene molecules first appeared to him in a *dream* (see [Hempel 1966](#) , 16). This fact was, of course, epistemically irrelevant and played no role in the justification of the hypothesis. By contrast, the *Boyle-Mariotte law of gases* was discovered by means of the systematic extrapolation of data ([Langley et al. 1987](#) , 81) – and this fact played a decisive role in the justification of the hypothesis.

Philosophers should *learn* from the de facto methods and arguments of science as far as they are successful, and they can only do this by observing them. Instead of premise 2, all the normativists usually want to say, or at least all they should say, is the following *almost trivial* weakening:

(2.1–2) *Premise 2** : The epistemically *external* part of the history of the development of a hypothesis is irrelevant to the assessment of its epistemic value.

Of course, the conclusion of argument ([2.1-1](#)) no longer follows if premise 2 is replaced by the weaker premise 2*. And with this conclusion, the error of the normativist argumentation becomes *grave* . To realize this, we must simply ask the following question: how can philosophers of science *acquire* the appropriate rules and criteria for the scientific method? It would be naïve to presume that these rules and criteria might be gained through pure logic or mere intuition; science is much too complex for that. That would be like somebody trying to write a bridge-building manual without having examined the de facto practice of bridge building and typical examples of existing bridges. The philosophy of science, then, *can* and *should* learn from the real examples provided by the history of science. This brings us to the following result: although it is the *primary* task of the philosophy of science to explain the methods and criteria of *rational* science, it can only fulfill this task by *also* studying the *factual sciences* in their historical development.

Kuhn's criticism crystallized into the insight that the scientific models provided by logical empiricism and critical rationalism were too *simplicistic* to give a faithful picture of real science (cf. [Stegmüller 1979b](#) , 5; 1969ff, vol. II.3, 20). Reacting to this, young theorists proclaimed that the philosophy of science should even be restricted to the descriptive analysis of the sciences. But this position too is overstated. The question of the definition and criteria of scientific rationality must, of course, remain central to philosophy of science, as it is this question which defines philosophy of science as a discipline and sets it apart from the history and sociology of science. In summary, the philosophy of science is a discipline, which contains both descriptive and normative elements – an insight that has been worked out by [Stegmüller \(1969](#) ff, vol. IV.1, 9ff).

The philosophy of science's method may best be described as a *rational reconstruction* . My understanding of this method is different from Lakatos' understanding of it, which has been sharply criticized by Laudan ([1977](#) , 167–170). For Lakatos ([1971](#) , 119f), 'history' as understood by rational reconstruction is not the *actual* history, but a projection of ideals of rationality into history, and a selection of those parts of history that conform to that ideal. By contrast, in my understanding, the actual history forms one pole of the method.

More precisely, rational reconstruction moves between two poles: a (so-called) *descriptive corrective* , which should be adequately reconstructed, and a *normative corrective* , which contains rational norms. The descriptive corrective does not contain every piece of factual sciences. It contains only typical examples of de facto successful (i.e. unanimously and sustainably accepted) scientific discoveries. Moreover, it also contains typical 'counterexamples,' that is, examples of de facto refuted scientific hypotheses. Philosophy of science bases its reconstructions on these consensually accepted examples of scientific success or failure, though only in a preliminary and defeasible way.

The normative corrective consists of certain epistemic intuitions that are normally assumed by scientists, in spite of their various cognitive differences. That such a common ground exists is the position that will be defended in this chapter. First of all, the normative corrective contains one supreme epistemic goal or norm, which may be explicated as follows (cf. [Weingartner 1978](#) , [sec. 3.2](#)):

(Prop. 2.1-1) The *supreme epistemic goal* (G) of science is to find *true* and *content-rich* statements, laws, or theories, relating to a given domain of phenomena.

A sentence is the richer in content, the more *consequences* it has (see [sec. 3.7](#)). The restriction of the goal of science to true *and* content-rich statements is significant, because the probability and the content of hypotheses are often *inversely proportional*. We can maximize the truth chances of hypotheses by sticking to trivial tautologies like: “The Sun revolves around the Earth, unless it doesn’t.” Inversely, it is easy to present very substantial and impressive hypotheses if we disregard their truth value, for example, “I have invented a perpetual motion machine.” The true art of the scientist is to formulate hypotheses which have both empirical truth and substantial content and consequences.

In the above explication of the supreme goal of science, there appears the little word “truth.” One might object that the goal of truth is not metaphysically neutral. However, we understand this concept in a liberal way, leaving room for different metaphysical interpretations (see [sec. 2.8.1](#)). On the other hand, we don’t want to be too liberal about the meaning of “truth.” For people have rather different ideas of what truth is. While one person seeks truth in the platonic heaven, another person may equate truth with empirical success, and for a third one truth is what makes herself happy. To ensure that the goal of scientific discovery as explicated in [prop. 2.1-1](#) is not entirely vague concerning the meaning of “truth,” we need a *minimal* explication of what “search for content-rich truths” is intended to mean. This is provided by the second ‘normative’ component that is normally assumed by scientists as well as by philosophers of science, namely a minimal common epistemological model of scientific discovery. In [sec. 2.2](#), we will explain this minimal epistemological model of the sciences in detail. The key points are given here: it contains (1) a *minimal realism*, (2) *fallibilism* and critical attitude, (3) the struggle for *objectivity*, (4) a *minimal empiricism*, and (5) the struggle for *logical clarity*. Although this minimal epistemological model consists primarily not of normative, but of *descriptive* claims, we still regard it as part of the normative corrective of philosophy of science, as it gives clear meaning to the supreme goal of science, and is usually seen as a precondition of science.

Given the above described normative and descriptive correctives, the method of rational reconstruction may be precisely characterized as follows:

(Def. 2.1-1) *Method of rational reconstruction* : Develop generalized and logically precise *models* of scientific discovery, which on the one hand *fit the descriptive corrective*, that is, provide optimal explanations for the success of the examples and the failure of the counterexamples from the history of science, and which on the other hand may be *justified according to the normative corrective*, as providing optimal *means* of reaching the supreme goal of scientific discovery, given the assumptions of the minimal epistemological model.

In short, the philosophy of science abstracts generalized models from the descriptive corrective, which at the same time are (or should be) in line with the normative corrective. This is summarized in the schematic diagram in [fig. 2.1-1](#).

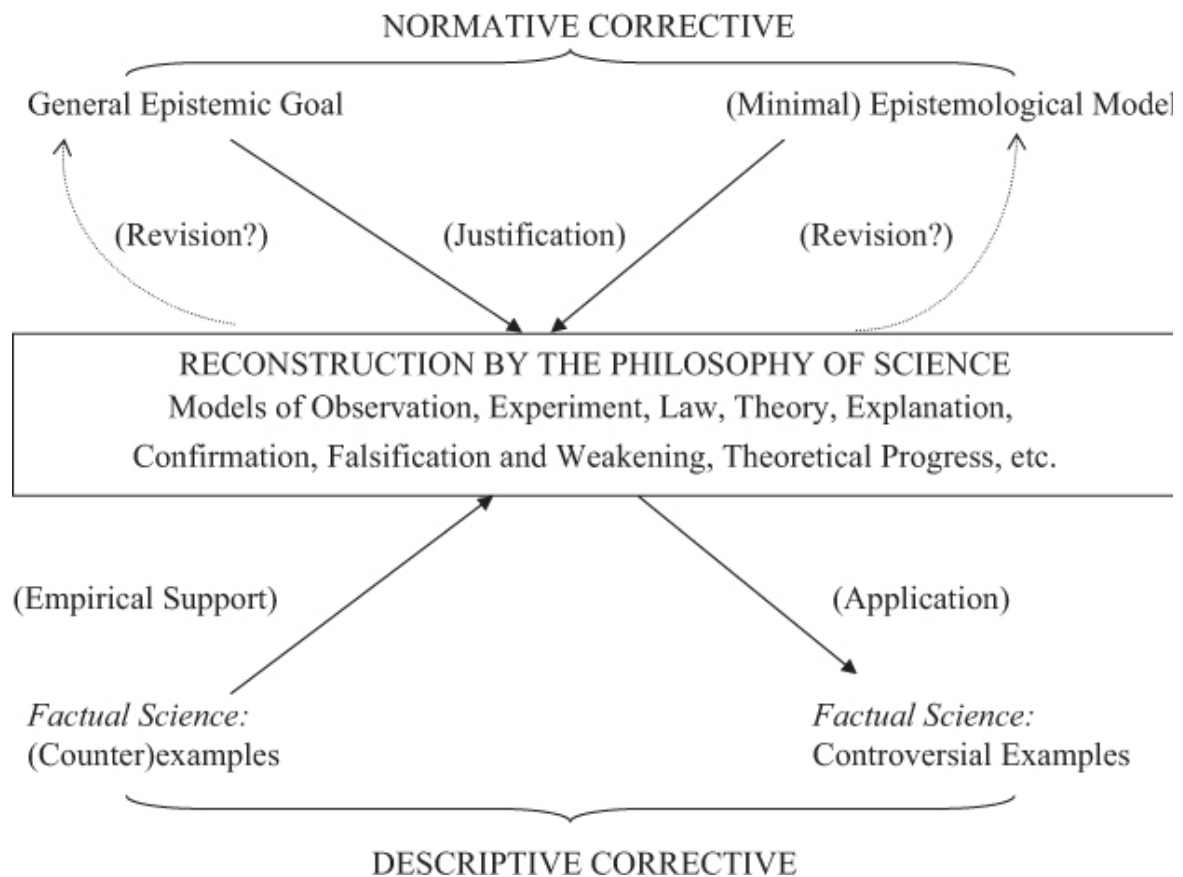


Figure 2.1-1 The method of rational reconstruction. [↗](#)

As long as both aspects of rational reconstruction – descriptive adequacy and normative justifiability – can be made to work in harmony with each other, the work of scientific reconstruction can be counted as successful and its results as *corroborated*. Of course, we cannot expect that everything we find in *factual science* will also be *rational* according to the normative corrective. For this reason, the descriptive corrective of the philosophy of science contains only *accepted* (counter) examples. When models developed by the philosophy of science have been corroborated, we can take the next step of applying them to *controversial* areas of scientific disciplines, in which no consensus has been reached – be it the objectivity problem of quantum mechanics, the teleology problem in biology, or the explanation-understanding problem in the humanities. It is exactly here that philosophy of science can supply valuable decision-making aids to the individual sciences. So, on the one hand, the philosophy of science learns from the individual sciences and is corrected by them; on the other hand, its reconstruction work can be instructive to the individual sciences and thus help them solve specific problems.

The goal of rational reconstruction, though, bringing together descriptive analysis and normative justification, does not always have to succeed. This applies, for example, when philosophers of science proceed from unrealistically high or one-sided normative standards, which will persistently disallow them to reconcile their normative corrective with the descriptive analysis of science. In such cases, the philosophy of science must go on to *revise* its own normative corrective. When it does this – this has already happened several times in the history of the philosophy of science – then the philosophy of science, to borrow Kuhn's terminology, is no longer in a normal scientific phase, but in a revolutionary phase.

To conclude this section, we will compare the model of rational reconstruction with two similar characterizations of the method of philosophy of science. Carnap (1950b, sec. 2) described the procedure of the philosophy of science as *explication*. Here, a vague concept of everyday or scientific language (the explicandum), for example, the idea of probability, is replaced by a precisely defined term (the explicatum), whereby the explicatum must differ from the explicandum insofar as it eliminates vagueness and ambiguity,

but it should also remain as close to it as it can and be as fruitful and simple as possible. Brun ([2016](#)) describes Carnap's method of explication as "conceptual re-engineering." Carnap does not explain, however, how exactly these partially contradictory demands should be combined. In comparison, the proposed model of rational reconstruction contains clear normative and descriptive guidelines; and by distinguishing normal and revolutionary phases in the philosophy of science, it takes into account the possibility that incoherencies can arise between the two. There are also some similarities between rational reconstruction and the method of reflective equilibrium according to Goodman ([1955](#) , 85–89) and Rawls ([1971](#) , 20f, 48–52). However, while the method of reflective equilibrium works purely *coherently* , aiming at the reciprocal adaptation of methodological rules and intuition, rational reconstruction is *empirically grounded* by its descriptive corrective. Moreover, the normative corrective consists not of mere intuition, but of epistemological *hypotheses* , together with a supreme epistemological norm.

2.2 Common Epistemological Assumptions

The minimal epistemological model, common to more or less all scientific disciplines, contains five epistemological assumptions (E1–E5):

E1 – Minimal realism : According to this assumption, there is a reality, which exists independently of the (given) epistemological subject. It is not assumed that all properties of this reality are knowable. The possibility of fundamental limitations to scientific knowledge is left open and cannot be answered a priori, but only based on the factual epistemic success of science. Scientific disciplines aim to provide statements which are as true and as content-rich as possible about clearly defined areas of this reality. The term "truth" is used here in the sense of the *structural correspondence theory* , according to which the truth of a sentence consists in its structural agreement with the part of reality described by it. This notion of structural correspondence, which was first precisely defined by Alfred Tarski ([1936a](#)), does *not* imply any intrinsic relationship or mirroring between language and reality. It also allows that theories which are strictly speaking *false* may be true in an *approximative* sense, or *close* to the truth ([sec. 5.6.2](#) , 5.11.1).

E2 – Fallibilism and critical attitude : There is no infallible 'golden path' to the truth in the sense of the correspondence theory. According to the assumption of fallibilism, every scientific statement is more or less fallible; so, we can never be *absolutely* sure of their truth. Fallibility goes hand in hand with a critical attitude, according to which no statement may ever be once and for all exempted from criticism. However, we can regard the truth of a statement as more or less probable, and in the limiting case as "practically certain," in which case we consider its error probability negligible. For this reason, everything depends on using empirical tests to find out more about the *probability* of a scientific hypothesis, in the sense of assumption E4 below.

E3 – Objectivity and intersubjectivity : According to this assumption, a statement's truth must hold objectively, that is, independently of the beliefs and value-attitudes of the epistemological subject, because by assumption E1, reality has subject-independent existence, and truth is the correspondence between statement and reality. So, the assumption of objectivity already follows from assumption E1. However, the characterization of objectivity as subject-independence does not help us in the practice of science, as only subjects make statements and formulate hypotheses. We have no direct access to objective truth; we can only approach it indirectly via – not certain, but probable – *criteria* .

The second part of assumption E3 says that a central criterion for objectivity and also for truth lies in the *intersubjectivity* of scientific statements. Intersubjectivity, correctly understood, is a plausible consequence as well as a plausible indicator of objectivity and truth. It is a plausible consequence because if a statement's truth can be established at all convincingly, then – at least in principle – it must be possible to convince any person of the statement's truth, provided the person is sufficiently competent, that is, cognitively trained and informed. And it is a plausible indicator, because the agreement of many mutually *independent* and competent (i.e. fallible but truth-biased) researchers on the hypothesis provably increases its truth chances. This is an instance of the *Condorcet jury theorem* (cf. [Bovens/Hartmann 2003](#) , [sec. 5.2](#) ; [Dietrich/Spiekermann 2023](#)).

E4 – Minimal empiricism : With the exception of formal sciences such as mathematics, which we will discuss later, the objects examined by any science must in principle be accessible to experience and observation. After all, reliable information about reality can only be gained through perceptual observation; for perception is our only form of direct informational contact with reality. Empirical observations, then, are a decisive referee in the scientific search for truth: scientific law hypotheses and theories must be *tested* in terms of them. This is the central statement of assumption E4. Observations are not infallible (that would contradict assumption E2), but they offer the easiest and fastest route to intersubjectivity and practical certainty. Observations are “epistemically privileged,” as one also says. The empiricism contained in assumption E4 is *minimal*, because different from empirical reductionism it is not claimed that all scientific statements must be traceable back to observations via a chain of definitions, or must be provable by observations. Scientific theories can and should contain theoretical concepts, by which they can talk about things which lie beyond what is observable to human senses. The decisive requirement of E4 is merely that statements about the unobservable must have *observable consequences* by which they may be tested.

E5 – Logic in the wider sense : Using precise logical methods to introduce terms, to formulate sentences, and to construct correct arguments, is the most effective way to approach the goal of content-rich truth (according to assumptions E1–E4). For the meaning of a sentence cannot be established until the terms used in it have been clearly defined. And only sentences with a precisely formulated meaning may be examined for their logical consequences. Finally, only after the consequences of a hypothesis are known, may this hypothesis be precisely empirically tested in accordance with assumption E4. In conclusion, the procedure of empirical testing requires the use of exact logical methods at all times – in the wider sense of logic, not restricted to deductive logic.

Our minimal epistemological model makes as few assumptions as possible, but as many as necessary, in order to explain the possibility of objective science. Some of its assumptions are worthy of further explanation. We call the realism in assumption E1 “minimal” because it only asserts the objective *existence* of reality. Whether and to what extent reality is objectively discoverable cannot, according to assumption E2, be answered a priori, but only through the success of scientific discovery methods. This sets our minimal realism apart from metaphysical realism, in all its varieties. Moreover, we don’t claim that science is impossible without the assumption of a subject-independent reality, but only that this assumption is *normally* made in scientific disciplines (see [sec. 2.8.1](#)).

To avoid misunderstandings, we should add that the correspondence theory of truth, while well suited as a *definition* of truth, does not offer any *criteria* to evaluate the truth of a sentence (cf. [Rescher 1973](#) , ch. 1; [Kirkham 1992](#) , 195). The path to such criteria is illuminated by assumptions E5 and E4: investigate the empirical consequences of your hypothesis and test them using empirical observations. For example, to find out the truth of the sentence: “this flower is red,” the correspondence-theoretical insight that this sentence is true if the flower is red will not get me anywhere; rather, I must be in a position to apprehend the object denoted by the ostensive individual term “this flower” and the perceptual quality expressed by the predicate “red” visually, in order to test the truth of the sentence by a simple act of perception.

In assumptions E4 and E5, we claim the existence of only two minimal levels, in which discovery finds its origin and its touchstone: observation, which allows us fragmentary but reliable and intersubjective access to reality, and logic in its wider sense, which provides a universal cognitive tool, but in itself says nothing about reality.

In assumption E3, we limited our characterization of *intersubjectivity* to “sufficiently competent and informed” persons. This limitation is necessary, because persons who do not *understand* the content of a hypothesis, or who do not possess the relevant *evidence* , are unable to test a hypothesis for its truth. This difficulty shows us that intersubjectivity can on no account be used as a *definition* of objectivity or truth. The latter suggestion has been made in the constructivist school of Kamlah and Lorenzen ([1984](#) , 119), and one century earlier by the founder of the philosophical school of *pragmati(ci)sm* , Charles S. Peirce ([1878b](#)). But competence is an uncertain and gradual criterion, meaning that any number of persons judged to be competent may still err collectively. Peirce attempted to save his consensus definition by equating truth with the consensus that an *ideal* community of researchers with an infinite amount of time would approach in the limit. But even in this counterfactual version, an intersubjective consensus does not have to agree with the

truth of correspondence theory. There are countless questions on whose true answer even an ideal community of researchers can never achieve consensus, because there is no evidential access to them, although there is no reason to assume that these propositions have an indefinite truth value – for example, “How many microbes were located in Caesar’s nose as he crossed the Rubicon?” In conclusion, intersubjectivity can at most be an uncertain criterion for objectivity.

2.3 Common Methodological Features

From the supreme epistemic goal (G) and the five epistemological assumptions (E1–5), there follow very plausibly four methodological features (M1–4), which are common to all empirical sciences in the broad sense. First, a word on terminology. By an *observation sentence*, we understand a sentence that expresses an *observable* state of affairs. We distinguish between *actual* and *potential* observation sentences. We call an observation sentence actual if the state of affairs expressed by it has indeed *been observed*. By contrast, an observation sentence is merely potential, if the state of affairs expressed by it has not been observed, but could be observed, for example, at a later time or in a different place. Merely potential observation sentences express empirical *predictions*, for example, “there will be a solar eclipse tomorrow.” We can now formulate the four common methodological features as follows (similarly in [Weingartner 1978](#), [sec. 3.1](#) – [3.7](#)):

- M1* : Science searches for hypotheses which are as general and as content-rich as possible and recorded in a scientific language. In all disciplines, these hypotheses include *laws* and *theories*, and in some (e.g. the historical sciences), they also include *hypothetical singular sentences*.
- M2* : Science searches for *actual observation sentences*, as many as possible (and as relevant as possible), which reflect the results of observations, experiments, or measurements.
- M3* : Science attempts, with the help of general and hypothetical sentences, to *explain* the currently known actual observation sentences, and to *predict* new, and as yet unknown, *potential* observation sentences.
- M4* : Science attempts to *test empirically* its general and hypothetical sentences by comparing the predicted (potential) observation sentences with the currently known (actual) observation sentences. If the latter are in agreement with the former, the prediction was successful (it then becomes a successful explanation) and the hypothetical sentence (law or theory) is *confirmed* or *corroborated*. If there is a contradiction between the latter and the former, the prediction was without success and the law or theory is *falsified*, or in the case of a merely statistical prediction, *weakened*.

We only speak of an *explanation* if the state of affairs to be explained has already come to pass and is known, while in a *prediction*, the predicted state of affairs is new and unknown. Viewed logically, both explanations and predictions take the form of deductive or probabilistic arguments (see [sec. 6.1](#) – [6.4](#)).

These four methodological features or modules are closely connected with the epistemological assumptions of science. Feature M1 results from the epistemic goal G, because laws and theories are the typical examples of general, content-rich sentences. M2 results from assumption E4 (sharpened by assumptions E2 and E3): to discover more about the conjectured truth of its hypotheses, science must collect empirical data. In order for general hypotheses to be tested on the basis of these data, the logical or probabilistic consequences of the hypotheses, in the form of predictions or potential explanations, must be investigated, and these consequences compared with the data – in this way, the methodological building blocks M3 and M4 are motivated by the assumptions E4 and E5.

Successful predictions and explanations not only have a bearing on the ultimate scientific goal of finding true and content-rich theories, but they are also of crucial importance for human *planning behavior*, indeed, to our survival. Predictions allow people to plan their future by *adapting themselves* to the predicted sequence of events. Explanations are concerned with finding causes ([sec. 6.2](#)); they provide people with *knowledge of causes*. Causal knowledge gives people more power than merely predictive knowledge: it enables them to causally intervene in the course of events, and thereby *adapt* the world to suit *their own wishes*. The scientific subgoals of prediction and causal control have been pursued by humanity since the Stone Age. Anyone who can *predict* that nourishing plants are growing in a given area can decide to migrate there and thereby ensure a high reproduction rate for their own tribal group. But those who know the *causes* of the growth of nourishing plants, who know how to plant and cultivate them, need no longer be gatherers, but may become farmers, and practice agriculture instead.

In advanced science, the scientific procedures of explanation and prediction operate not on two but on three levels: actual observation sentences at the lowest level, empirical law hypotheses at the middle level, and scientific theories at the upper level. Between the first and the second, and between the second and the third level, the methodological procedures from M3 and M4 take place, as shown in [fig. 2.3-1](#). This threefold division of scientific work is advantageous for many reasons; in all advanced scientific disciplines, a division of labor has arisen between practical experimenters (levels 1 and 2) and theoreticians (levels 2 and 3). To give an example: In the case of the physical theory of the planetary orbits ([sec. 5.1](#)), level 1 corresponds to the actual astronomic observations (e.g. by Tycho Brahe), level 2 corresponds to the empirical laws of the (elliptical) planetary orbits around the Sun (inductively obtained by Johannes Kepler), and level 3 to Isaac Newton's theoretical explanations of these orbits by the Sun's gravitational attraction together with the planets' centrifugal force.

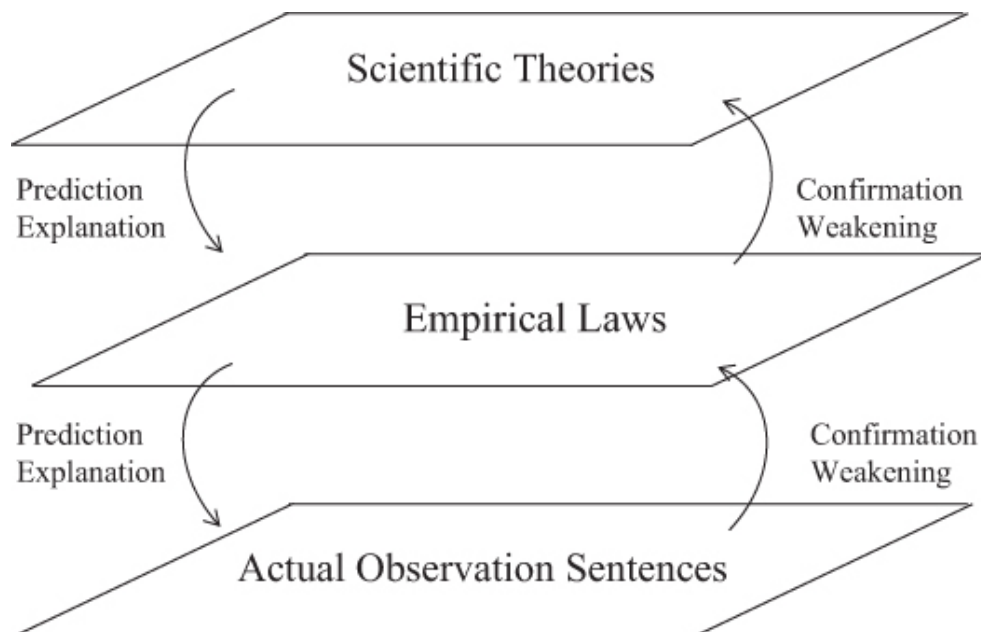


Figure 2.3-1: The three levels of the scientific method. [↗](#)

2.4 Scientific Disciplines and Their Classification

In this chapter, we will examine to what extent our epistemological assumptions E1–5 and the methodological features M1–4 match the existing scientific disciplines. To structure this investigation systematically, we introduce something that is rarely found in textbooks on the philosophy of science: a division of scientific disciplines according to their area of inquiry (see 2.4-1). Our division is certainly imperfect and we attribute to it not a final but merely a heuristic status. Some disciplines are mixed disciplines and are listed under “also;” further special disciplines can be added:

(2.4-1): *Scientific disciplines, classified according to area of inquiry :*

Sciences...

1. ... *of nature* : physics, chemistry, biology, geology, medicine (astronomy, cosmology, geography, paleontology, history of biological evolution)
2. ... *of technology* : mechanical and electrical engineering ..., also: computer science
3. ... *of human beings* : psychology (also: education, medicine, cognitive science)
4. ... *of (human) society* : sociology, economics, political science (also: anthropology, ethnology, geography)
5. ... *of (human) history* : history (also: anthropology, ethnology, philosophy as history of ideas)

6. ... *of human cultural (mental, social) artifacts* : legal sciences, linguistics, literary science, sciences of fine arts and music, art history, media studies (also: education, religious studies)
7. ... *of formal structures (formal sciences)* : mathematics (logic, statistics, theoretical computer science, system theory...), formal methodology and philosophy of science
8. ... *of the general foundations of human ideas* : philosophy (epistemology, philosophy of science and theoretical philosophy; ethics, esthetics, and practical philosophy)
9. ... *of God* : theology (also: religious studies)

Our classification has the advantage of not being tainted by partly value-based categories such as (natural) *science*, *humanities*, or *real science* (etc.). If we try to apply these catch-all categories to our area-of-inquiry-based classification, we recognize that there is no clear consensus about what goes where, and that different systems of classification often have more to do with the politics than the philosophy of science. Group 1, for example, covers the core *natural sciences*. But mathematics too sees itself as a natural science, although it belongs to the category of *formal sciences*; moreover, in the meantime, mathematics is not only applied in natural sciences but also in economics and social sciences. Of late, psychology also sees itself predominantly as a natural science, although it could equally be assigned to the humanities or social sciences.

Groups 5 and 6 comprise the core of the traditional *humanities*. Philosophy, though, is just as close to the natural sciences as it is to the humanities, although in contemporary universities, it is usually placed in the faculty of arts and humanities. In [sec. 1.3.8 – 1.3.9](#), we explained that since the 19th century, the traditional humanities were increasingly understood as a counterpart or even an opposition to the natural sciences. The term *humanities and social sciences* was coined in the 1960s and intended to serve as a modernized counterpart to the natural sciences. Whether this terminological group covers only groups 3 and 4, or also groups 5 and 6, was not terminologically fixed.

In the last two decades, the new terminological group of the *cultural sciences* was formed, in part as a reaction to contemporary tendencies in psychology and social sciences to increasingly use methods of natural sciences. The term “cultural sciences” is predominantly used in continental Europe and is understood in a broader sense than the Anglo-Saxon “cultural studies,” though its extension is not precisely defined. The frequently supposed opposition between natural and cultural sciences is an instance of what Snow ([1960](#)) has called the two ‘cultures’ in the current scientific age. In this book, we intend to refute the view that there exists a principal methodological *gap* between the sciences of nature and culture; this gap, where it exists, is more a matter of different value-attitudes and worldviews than a matter of unbridgeable differences in method.

In the term *factual sciences* (in German “Realwissenschaft”), all sciences are brought together that have a part of the *real* world as an area of inquiry; so, factual sciences include all groups apart from the formal sciences. Theology can only be seen as a factual science if we assume the reality of God, as this reality cannot be scientifically demonstrated, its inclusion here is questionable.

2.4.1 The Special Status of Mathematical Sciences

First, we have to explain the special status of group 7, mathematics, and formal sciences. In their totality, assumptions E1–5 and features M1–4 only make sense for the factual-empirical sciences. Mathematics and formal sciences do not have a domain of inquiry that consists of *concrete* and empirically accessible objects, but rather deal with formal and *abstract* structures. Mathematics, then, has no empirical basis, and therefore assumption E5 and methodological features M2 and M4, which require an empirical basis, do not apply. This does not mean that mathematics is a speculative discipline, which makes statements about reality with no empirical basis: pure mathematics makes no statements at all about an (empirically accessible) *reality*. All it does is develop formal structures and models, characterize them by using abstract axioms and definitions which are as general as possible, and derive logical consequences from them. It would be nonsense, for example, to say that the axiomatic system of real numbers or of Euclidean geometry was per se “true” or “false.” Such claims can only be made when mathematical terms are interpreted empirically, for example, when real locations or time points are assigned to real numbers, or real distances to mathematical vectors. But then geometry is no longer a mathematical theory, but becomes a physical one. Einstein once formulated it like this: “As far as the theorems of mathematics relate to reality, they are not certain; and as far as they are certain, they do not relate to reality” (cf. [Carnap 1967](#), 56). So, we can say: exactly *because* mathematics

ignores the empirical interpretation of its models, it is the strictest of all sciences. To be sure, mathematics concerns itself primarily with those structural models which are applicable to certain areas of reality. Seen in this way, mathematical disciplines can be more or less *fruitful* for the real sciences, but they cannot be true or false in the sense of the factual sciences.

Assumptions E2–4 and the methodological feature M1 can be applied to mathematics without restriction (the assumption of fallibility E2 insofar as even mathematical proofs are not guaranteed to be error-free). Assumption E1 of minimal realism applies only in an indirect way, insofar as mathematical structures have the purpose of being applied to the factual sciences. The methodological component M3 applies insofar as mathematicians are interested in the derivation of the various theorems, which express important properties of their models. Such derivations may be termed mathematical *explanations*, as long as it is clear that these are not causal explanations of real facts (cf. [Weingartner 1978](#), 102; [Pincock 2015](#); [Kuorikoski 2021](#)).

Let us not neglect the fact that the *philosophy of mathematics* includes opposing views on the nature of mathematical objects (cf. [Benacerraf/Putnam 1964](#); [Shapiro 2000](#); [Linnebo 2020](#)). According to the widespread view which is shared by us, mathematical objects are abstract *conceptual* objects, that is, conceptual constructions. Real objects, for example, five red roses, can *exemplify* mathematical objects, for example, the number five, but can never be identical with them. This view, called *conceptual structuralism*, was worked out in detail by the group of mathematicians around Bourbaki ([1961](#)) and philosophers of mathematics like Field ([1980](#)). The opposing standpoint in the philosophy of mathematics is *mathematical realism*, which assumes that mathematical objects exist *per se*. This position has been defended, for example, by Frege ([1918](#), 43) and Gödel ([1947](#), 271f). Also in this position, mathematical entities are seen as abstract and *platonic* objects, not located in the reality of space and time, but in a “platonic world of mathematical ideas” which is accessible not empirically, but only through mathematical “intuition.” An argument in favor of mathematical structuralism is Ockham’s razor (cf. 5.8.7): structuralism explains mathematical phenomena equally well as realism, but without the assumption of strange platonic entities. Another position is mathematical naturalism ([Maddy 1997](#)), according to which mathematical theories resemble scientific theories because they can fit reality more or less well in their applications. (The analogous view in the philosophy of logic has been called “anti-exceptionalism;” cf. [Hjortland 2017](#).) Against this view, it must be objected that in mathematical theories the reference to factual applications is fundamentally left open, whereas in empirical theories this reference is semantically fixed. As a result, empirical theories can be tested through experiments, which is undoubtedly *not* the case for mathematical theories.

In summary, mathematics and the other formal sciences are a *specialization* within the whole system of science, dealing with formal models, algorithms, and methods. This specialization has shown itself to be particularly *useful* in the history of science. Put in a nutshell, the information that the mathematician passes to the empirical scientist is this: “*if* your empirical domain satisfies the axioms of this-and-this mathematical structure, *then* it must also satisfy these-and-these theorems about this structure that we have proved; we grant you the satisfaction of the theorems, provided *you* grant us the satisfaction of the axioms.”

So, we characterize the formal sciences as *prerequisite and auxiliary sciences*, which furnish the factual sciences with formal foundations and structural models with known properties.

2.4.2 Testing the Common Foundations of Science

Having clarified the special status of mathematics, let us now ask: with the *exception* of the formal sciences, to what degree do the assumptions E1–5 and the methodological features M1–4 comply with the listed groups of disciplines? Obviously, E1–5 and M1–4 fit all disciplines in groups 1–3 (natural sciences, technical disciplines, psychology). There is also no reason not to apply them to the social sciences and historical disciplines in groups 4 and 5. Minimal realism, fallibilism, objectivity and intersubjectivity (E1–3), and logical clarity (E5) are accepted standards in all these disciplines, and the same is true for the search for general hypotheses, explanations, and predictions (M1, M3). In all these disciplines, empirical data are available, for example, in the form of the results of psychological experiments, interview data, economic and demographic data, and in the historical sciences in the form of historical texts and source materials, etc. By this, assumption E4 and building blocks M2 and M4 are also fulfilled. Of course, it is significantly more difficult to find general law hypotheses in social science and history, because of the complexity of their area of inquiry – but all that is required for the applicability of the methodological building blocks to a given

science is that the goals therein can reasonably be *sought after*, which does not necessarily mean that these goals are equally easy to attain in all different disciplines.

The same considerations apply to broad areas of the disciplines in group 6. Here as well, the epistemological assumptions E1–3 and (to a lesser extent) E5 are widely accepted standards. Moreover, these disciplines work with empirical data too; so, assumption E4 and the methodological building block M2 apply. For example, linguistic texts comprise the data for linguistics, and works of art and music from various cultures or stylistic periods are the data for literary criticism, art history, and music studies ([Kreuz/MacNealy 1996](#)). Members of the hermeneutic school raise the objection at this point, that “data” in the social sciences and humanities, for example, speech acts or texts, are always of an *interpretative* nature, so that understanding in the humanities is always subject to an unbreakable *hermeneutic circle* (cf. [Gadamer 1975](#) /2005, 235ff). But a number of philosophers of science have worked on the view that the testing of interpretative hypotheses concerning what an author intended to say proceeds in the same “holistic” manner as the testing of theories in the natural sciences (cf. Stegmüller [1979a](#) , 27–86; [Follesdal 1979](#) ; [Rosenberg 2008](#) , 43f). What hermeneutic philosophers rightly point out is that what one calls ‘data’ in the humanities are usually not raw data, but semantically interpreted data in the form of meaningful linguistic or nonlinguistic acts. However, also what physicists call data are usually theoretically interpreted measurement results, for example, elementary particles observed in the cloud chamber. This does *not* mean that raw or uninterpreted data or stimuli do not exist at all: in our physical example, they exist as condensation droplets in the cloud chamber; and in the interpretation of linguistic acts, they exist in the form of sequences of sounds or letters. Although some advocates of hermeneutics doubt the existence of raw linguistic data, their existence can easily be verified: you just have to hear someone speaking in a completely foreign language; what you will then hear are exactly these raw linguistic data.

The school of *methodological dualism* mentioned in [sec. 1.3.8](#) objects, once more, that there are no general law hypotheses in the humanities. If that were true, then the activities of these disciplines would have to be limited to methodological building block M2, that is to say, to the gathering of singular historical facts according to the “ideographic method.” What is true is that in the humanities and the social sciences, there are as good as no *strict* (or exceptionless) laws. But, as will be worked out later ([sec. 4.5](#) , [6.4.2](#)), there are *statistical* and in particular *normic* law hypotheses (“if A, then normally B”), which are indispensable for explanations and hypothesis testing in these disciplines. Therefore, the methodological components M1, M3, and M4 apply also to the historical sciences and humanities.

To avoid misunderstandings, when we say that the methodological components M1–M4 can be and have been successfully applied in the humanities, we don’t imply that *all* representatives of this group of disciplines make use of these methods. Within the humanities and social sciences, there is still the opposition between proponents and opponents of empirical-analytic methods – we mentioned Snow’s (1960) “two cultures” and Ross’s (1996) thesis of the “science wars” ([sec. 1.3.9](#)). Some scientists in the humanities and social sciences reject these methods simply *because* they are empirical and analytic. What this book nevertheless intends to show is that the methodological features M1–M4 are optimal means to pursue the supreme goal (G) of finding content-rich truths even in the human and social sciences.

There is, however, another phenomenon, found mainly in the humanities, that seems to stand in decisive opposition to the scientific method according to E1–5 and M1–4: namely, the phenomenon of *norms and value judgments* . We illustrate this problem at hand of a subdiscipline of legal sciences: jurisprudence, that is, the science of legal judgment. This discipline examines concrete actions and the facts of a case to determine how they fit into the given legal system, and specifically whether they fall under a certain category of offense or crime (going along with a certain punishment). The philosophy of *legal positivism* , which goes back to [Kelsen \(1960/2002\)](#), held the view that jurisprudence is a value-free systematic-deductive procedure, based only on *positive* law, that is, laws which have been passed by legislators. But Kelsen’s thesis is controversial in the philosophy of law. For example, was the removal of private property from a grocer’s shop *petty larceny* or *theft* ? (an example from [Savigny et al. 1976](#) , 101f). Or: does an underwear advertisement poster still conform to ‘public decency,’ or does it already violate ‘the dignity of women?’ (an example from [Hilgendorf 2000](#) , 15, 24). In these and other cases, it does not follow from the stipulations of the law and empirical knowledge of the facts, how the deed of the accused should be classified; and the judges and other interpreters of the law now have the task of *interpreting* the underdefined stipulations of that law. In [Savigny et al. \(1976\)](#) , 106, 144ff), it is shown that these interpretations are often based on *value assumptions* (or assumptions of justness), which cannot be found in legal texts, but rather come from an

‘intuitive legal consciousness.’ Now, a legal scientist can certainly show that *if* certain value assumptions are taken for granted, *then* a certain interpretation of the law is adequate. But how can she determine scientifically *which* value assumptions are justified? Can the legal scientist prove that a specified degree of misery of the accused person is sufficient for us to speak of ‘petty larceny’ instead of theft? Can she scientifically decide whether a billboard advertising underwear is still in accord with ‘common decency’ or is already an affront to the ‘dignity of women’? So the question which arises *not only* in jurisprudence, but also in other disciplines relating to the human or social sciences, is this: can fundamental value assumptions be scientifically justified?

Without further ado, the answer proposed by us is *no*. This does not mean that value- and norm-related issues do not occur in science. For one thing, the de facto values of persons can be *empirically* investigated. Moreover, specific value sentences can be logically deduced from *assumed* fundamental value sentences (and further premises). But a fundamental scientific justification of values is impossible. Our view here is concordant with the general consensus of (empirical) science and does not claim to be original; but our justification of this answer should contain some originality. For the value-neutrality requirement has been criticized so sharply by critics of empirical-analytical philosophy of science that it simply cannot nowadays be taken for granted, but needs to be worked through with great care. This task is addressed in [sec. 2.5](#).

2.4.3 Two Further Classifications of Sciences

The above subdivision of scientific disciplines was given according to area of inquiry. To end this section, we will explain two further possible systems of classification. One methodologically significant classification of the factual sciences follows the “invasiveness” of their empirical discovery methods. The classification is summarized in [fig. 2.4-1](#).

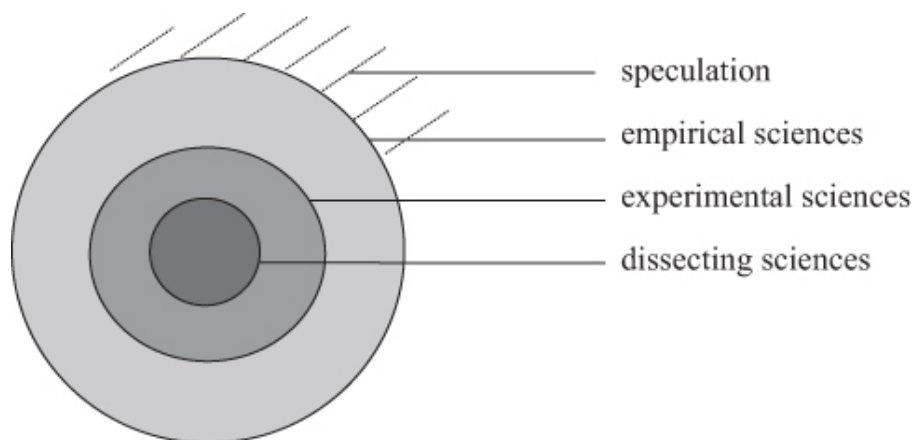


Figure 2.4-1 Classification of the factual sciences. [↗](#)

1. *Empirical sciences* (in the sense of minimal empiricism): By this, we understand all disciplines which refer in some way to empirical data, whether these data were obtained via free field observation, by studying historical records, or in laboratory experiments. To this most extensive group, we assign all disciplines described by assumptions E1–5 and methodological building blocks M1–4, including the humanities. Only those disciplines whose theories are based on pure *speculation* are not in this group – as for example mysticism, esotericism, speculative metaphysics, or the dogmatic and creed-based components of theology. Based on our epistemological assumptions, it is doubtful whether these are factual sciences at all.
2. *Experimental sciences*: This *subgroup* of empirical disciplines obtains its data from *controlled experiments*. The significance of controlled experiments as opposed to free field observation is that in an experiment, certain sets of conditions can be deliberately created, which give us information about the causal relevance or irrelevance of factors (see [sec. 4.2 – 4.4](#)), while field observers might have to wait ‘forever’ for a lucky coincidence to come up with the exact set of conditions required. For this reason, experimental sciences can explore their domain of inquiry much faster and more effectively

than sciences which rely on mere observation or passive reception of historical data. The experimental criterion separates the disciplines in groups 1–3 from the ones in groups 4–6, at least in tendency. For it is hardly possible, for both technical and ethical reasons, to perform experiments with large social groups, institutions, or legal systems. Of course, ethical concerns are already present in biological or psychological experiments, but here they are usually not strong enough to exclude them, only to restrict them.

3. *Dissecting sciences* : A significant subgroup of the experimental sciences are those which I will call – for want of a better word – the “dissecting” sciences. By this, I mean sciences, which not only expose their objects to various experimental conditions in order to analyze them intellectually, but which also *materially analyze* their objects, that is, take them apart or dissect them – perhaps in order to synthesize them afterward, that is, to put them back together again, provided this is still possible. I suggest that it is this dissecting character that constitutes the essential characteristic of *classical* natural sciences like physics, chemistry, biology, and medicine, in contrast say to experimental behavioral research or psychology.

Dissection of objects goes far beyond experimentation at the input-output level insofar as it makes the objects’ inner structure accessible to observation and experimentation. Chemistry and biochemistry have learned to break down items from our environment into their molecular and atomic constituents (analytical chemistry) and to put them back together again (synthetic chemistry). Nuclear physics has done the same at a *subatomic* level, here involving much greater amounts of energy. The analysis and synthesis of living creatures being worked on by science today – from genetic engineering and neuroengineering to artificial intelligence and robotics – could bring us to a point in the future where it is possible to genetically alter or even create a human or android being. Some people think these possibilities are terrifying; others that they are breathtaking. Here, as always, we should remember: scientific discoveries may be used for various purposes, both “good” and “evil.” But certainly, the possible risks and dangers going along with these technological possibilities are getting bigger and bigger. Therefore, the *ethics of science*, whose purpose is to set limits to the technological realization of the scientifically possible, is going to become more significant.

Independently of this, scientific disciplines may also be divided up according to the degree to which their concepts and *methodologies* are quantitative and mathematical, as follows:

Graduated division of scientific methodologies by increasing degree (←) of logical-mathematical and quantitative precision



Quantitative methods are on the left and *qualitative* methods on the right side of this spectrum. Note that on both sides we are looking at *scientific* methods – also qualitative hypotheses and methods can be pursued in a logically clean and empirically controlled way (cf. [sec. 3.1.4](#) ; [Denzin/Lincoln 2000](#)). The ideological polarization between quantitative and qualitative methods that is held by some qualitative researchers (e.g. [Lamnek 1988](#)) appears unnecessary and exaggerated. The strength of qualitative methods (e.g. case studies, narrative interviews) lies in advance of quantitative methods – in the exploration of relevant parameters and the generation of promising hypotheses. But a qualitative exploration has to be followed up by a quantitative-statistical analysis, as this is the only reliable way to test the generality of one’s hypothesis. That qualitative and quantitative methods are *complementary* in this sense is a widely held view among empirical researchers in the social sciences (cf. [Lazarsfeld/Barton 1951](#) ; [Giddens 1982](#) , 328); however, this view is not uncontroversial (cf. [Flyvbjerg 2006](#)).

2.5 The Question of Value Neutrality

2.5.1 Max Weber's Value-Freedom Postulate

The first detailed work on the value-neutrality requirement goes back to one of the main founders of sociology, namely, *Max Weber* (1864–1920). Weber was involved in a controversy with Gustav Schmoller, who propagated the view that social science should work out moral value judgments relating to a desired social order. Weber spoke out strongly against the pseudo-objectivity of such “podium values,” that is, value judgments presented with ‘scientific’ authority: Objectivity, said Weber, could only be reached in science if scientists limit themselves to factual, descriptive statements, and separate these clearly from their own values ([Weber 1922](#) /1949, ch. II and X).

The main reason for the self-imposed limitations of science to descriptive judgments, recognized by Weber, may be summarized thus: *Values are not properties inherent to objects themselves, but rather are based on subjective interpretations by us humans. Norms are also not objective facts, but man-made requirements, designed to aid the realization of certain values.* In the case of testing value and norm sentences, there is nothing which corresponds to an *empirical foundation* according to assumption E4 and methodological building blocks M2 and M4. At the same time, value and norm sentences are by no means purely formal claims, which might be justifiable by means of pure logic. There is therefore no objectivity in the sense of empirical science when it comes to norms and values. Ultimately, the *decision* for or against certain values or norms is a question of personal *freedom*.

This explanation of the value-freedom postulate is *prima facie* not only theoretically, but also practically plausible. Everyone knows from experience that finding a consensus on values between members of different cultures is much more difficult than finding a consensus on factual issues. Many strict Islamists would not take offense at Western science; the Western value system, however, seems twisted in their eyes. Nonetheless, the value-freedom requirement was massively criticized from the 1960s onward, especially by neo-Marxist movements such as the *critical theory*, and in the German-speaking area the Weber-Schmoller controversy flared up once more as the “positivism dispute” (cf. [Frisby 1972](#) ; [Adorno et al. 1976](#) ; [Dahms 1994](#)) and the “value judgment dispute” ([Albert/Topitsch 1971](#)).

Proponents of critical theory raised the objection that a purely descriptive or “positivist” method in social science had the effect of simply theoretically reproducing the status quo of oppressive social structures and thereby implicitly sanctioning them, while *critical science* must concern itself with *changing* oppressive social conditions ([Habermas 1966](#)).

This accusation is based on a fatal *misunderstanding* : descriptive research, of course, also includes research into natural laws, and thereby research into that which is naturally, technically, or practically *possible* and *impossible* . So, in fact, empirical social science is the very thing which can show us how social conditions may be *changed* , thereby contributing to the emancipation demanded by representatives of critical theory. In fact, almost all the “neopositivists” of the Vienna Circle had a *de facto* progressive political orientation and saw much of their work as being connected with social progress ([Dahms 1994](#) , 125; [Stadler 2001](#)). The only difference is that critical-empirical social science, as defended by Popper and Albert against Horkheimer and Adorno in the “positivism debate,” restricts itself to pointing out the *possibilities* for social change, and leaves normative decisions to the process of *democratic opinion-formation* , rather than wanting to justify these in a quasi-scientific way.

Even if the main accusation of critical theory was off the mark, the period that followed saw a number of legitimate objections being raised, which are in need of a more subtle discussion. Not only critical theorists (cf. [Dahms 1994](#) , 375), but also younger analytic philosophers warned that questions of values and norms should on no account be handed over to *irrationality* . Indeed, a rational treatment of value questions is both possible and necessary in society, and science has an important role to play in this regard. In this regard too, it is advisable to turn to Max Weber, who made subtle refinements to the value-neutrality requirement, in the face of just such objections:

First , according to Weber, it is absolutely possible for a social scientist to do research on the factual presence of value and norm systems, *without* passing judgment herself ([Weber 1922](#) /1949, 11f). To study polygamy in Islam, it is neither necessary to be an advocate nor to be an opponent of it. The statement that a

person or a society *believes* in certain values is not in itself a value statement, but a *descriptive* claim, whose confirmation requires only empirical data.

Second, also according to Weber ([1922](#) /1949, 21f), the scientist can identify *logical relationships* among value or norm sentences. Relevant to this task are the – meanwhile highly developed – disciplines of analytic ethics (e.g. [Frankena 1963](#)), deontic logic (e.g. [Aqvist 1984](#)), and decision theory (e.g. [Raiffa 1968](#)).

Third and most importantly, with the help of *means-end inferences*, the scientist can infer from *given* norms and descriptive knowledge a series of derived norms about optimal means of reaching accepted ends and pass these on as recommendations to society and politics. Because this represents the most important function of science for society, the next section is devoted to it.

2.5.2 Value Neutrality and Means-End Inferences

Finding suitable means for given ends is an important task of descriptive science. It obeys the following logical form:

(Def. 2.5-1) *The general scheme of means-end inferences* :

Descriptive means-end premise : M is in the given circumstances C a necessary – or alternatively an optimal* – means for the realization of end E.

Thus – hypothetical value- resp. norm conclusion : If end E is accepted as a fundamental value (or norm), then the realization of means M is a derived value (or norm).

Means-end inferences are accepted by most ethical theories as analytically *valid* (the asterisk* is explained below). The conclusion can be formulated for value statements (“P is or is not valuable”) and normative statements (“P ought or ought not to be done”). Both are analytically related, which can roughly be expressed by the formula: “The good ought to be done and the bad ought to be avoided” (for details, see [Schurz 1997](#) sec. 11.2–11.3). Value judgments cannot be analytically reduced to factual statements about the value sentiments of some or many people (as [Anderson 2004](#) proposed), but rather possess *normative force* .

The recommendation of means in the conclusion is not understood in a categorical, but in a hypothetical or conditional sense. That is, it is relativized to the assumed ‘fundamental’ or initial values (ends) E: *If* such-and-such fundamental values are assumed, *then* certain action recommendations follow as derived instrumental values. For conceptual clarification: A *categorical* value statement has the form of a basic value sentence “P is (or is not) valuable;” a *hypothetical* or conditional value statement takes the form of an if-then sentence “If P is (not) valuable, then Q is (not) valuable” (this distinction goes back to [Kant 1785](#)). A categorical value statement is *fundamental* if it cannot be derived from other categorical value statements using means-end inferences; otherwise, it is *derived* .

It is crucial for the value neutrality of science that recommendations of means are formulated hypothetically rather than categorically. Only in this way can a recommendation of means follow from factual knowledge and thus remain value-neutral itself, and only in this way can objectively value-neutral science be reconciled with societal value engagement. Value judgments derived hypothetically by science are also referred to as *instrumental* rationality; however, we avoid this term, as it implies a certain devaluation (“merely instrumental”), whereas what we mean is a genuine value engagement – only formulated hypothetically in order to clearly *separate* the factual component from the value component of the statement.

Another advantage is that, through the use of rich scientific knowledge, a large number of derived norms can be obtained from just a few fundamental norms. For example, the EU climate goal of becoming carbon-neutral by 2050 entails a multitude of recommendations of means. In this way, the burden of justification can be shifted – if not entirely, then to a large extent – from the normative to the descriptive.

In a means-end inference, it is assumed that M is either a necessary means – that is, every realization of the goal requires the realization of M – or an optimal means, meaning that among all practically possible alternatives, M maximizes total utility with regard to goal realization and side effects. Strictly speaking, the definition of an optimal means requires a listing and cost-benefit evaluation of all alternative means. In order for the hypothetical recommendation of optimal means to remain value-neutral, this cost-benefit evaluation must be included as a fundamental value assumption in the antecedent (the “if”-part) of the recommendation

– which is indicated by the asterisk. The precise reconstruction is therefore more complex than presented in [def. 2.5-1](#) (see [Schurz 2024b](#), ch. 3).

If M, on the other hand, is a merely *sufficient* means for an end E, then the means-end inference is not valid, because many sufficient means cause harm that outweighs the benefit of achieving the end. For example, to achieve the end of letting fresh air into a room, smashing a hole through the wall is a sufficient but not recommendable means; by contrast, opening a window is an optimal means, and a necessary means is some sort of opening to the outdoors.

Neither the inference to necessary nor to optimal means implies the mistaken thesis that “the end justifies the means:” if the costs of every possible realization of the end E exceed its benefits, then E should not be realized and can no longer be considered an acceptable fundamental value (given a consequentialist understanding of value).

The scientists or experts should take the fundamental norms from the knowledge *users* – that is, the public or their political representatives. Based on these norms, the scientists can then derive recommendations of means from scientific knowledge and pass them back to the knowledge users. Let us illustrate this by means of educational *teaching theories* (cf. [Flanders 1970](#) ; [Patry 1991](#)). Let us take it that *cognitive success* (i.e. mastery of material to be learned) is the fundamental goal of teaching. Based on empirical research, an education scientist may reach the descriptive hypothesis that cognitive success may be maximized if between 70% and 80% of teaching time is teacher-guided. The education scientist passes on this conclusion to teachers as a *derived* norm. This procedure is, however, only in accordance with the value-neutrality requirement if the means recommendation is formulated *hypothetically* : if your goal is to maximize cognitive success, then 70%–80% of class time should be teacher- guided.

If the explicit reference to assumed values is neglected, ideologically worrying consequences can result. Assume that in our example, the teachers have different fundamental goals than the scientists; they see *social learning* as an equally important goal to cognitive success. Let it be empirically shown, perhaps even by the same scientist, that the optimal means to reach this purpose is a maximum of just 60% of teacher-guided class time, alongside 40% of group activity. What if the scientist in our example, however, passes on her recommendation categorically rather than hypothetically, for example, she asserts *80% of class time should be teacher-guided* , and teachers believe him? Now the scientist is *manipulating* the teachers, because she is concealing the fact that her means recommendation is valid only for *her own* value assumptions. The teachers, who *do not* share these values, are being badly advised on how to reach their own chosen ends.

The same is true of many other situations, from medicine (a doctor advises a patient) to law (a lawyer advises a married couple about to get divorced). It is highly important in all such cases that the scientific “expert” makes the value assumptions *explicit* that she assumes in her means recommendations, as this is the only way that the “customer” being advised can find out whether these correspond to his own values. With these examples, we *emphatically* point out that the value-neutrality requirement does not have a dogmatic, but rather a *democratic* and *emancipatory* function. In [sec. 2.10](#) , we will defend this thesis in the context of contemporary debates on values and their role in science.

In summary, the central conclusion of this section is that value neutrality and value engagement in science are indeed compatible. The oversimplification of the value-neutrality requirement into the “value-free doctrine,” according to which values have “no place” in the empirical sciences, is therefore a serious misrepresentation that often serves as an easily criticized straw man for opponents of value neutrality. Sciences can and should use means-end reasoning to derive values from externally given fundamental values and pass them on to knowledge users in the form of hypothetical recommendations of means. Only the *justification of fundamental values* lies outside the scope of objective science.

2.5.3 Explication of the Value-Neutrality Requirement

To reach the final formulation of our value-neutrality requirement, we need to take two further steps of clarification.

1. Many authors have argued that science cannot be value free in an absolute sense, because the whole enterprise of “science” itself rests on certain goals and values: the so-called *science-internal* or epistemic values (e.g. [Schmidt 1959](#) , 648ff; [Doppelt 2007](#) , 190ff; [Sober 2007](#) , 110). In [sec. 2.1](#) , we identified the search for general and content-rich truths as the supreme epistemic goal of all sciences.

This fundamental internal value of science is accepted by all scientific disciplines, from which a range of derived epistemic values follow as means to achieve this supreme epistemic value (see [Doppelt 2007](#)) – in particular, the four methodological characteristics mentioned in [sec. 2.3](#). The value-neutrality thesis must therefore be restricted to fundamental *science-external* or non-epistemic value assumptions.

2. Is the restriction to fundamental science-external values and norms enough to give us a sustainable version of the value-neutrality requirement? *No* – insofar as one views science not merely as a process of justification, but as a *social process*. The process of scientific discovery is usually divided into three phases: the *context of origination* (CO) (also called “context of discovery”), in which the scientifically relevant problems are first identified and defined, the *context of justification* (CJ), in which data are gathered and systems of hypotheses generated and tested, and the *context of application* (CA), in which well-tested discoveries are used for various different social purposes. In both the CO and the CA, non-epistemic values inevitably play a role. Decisions about which research questions are considered important enough to pursue are also influenced by non-epistemic interests – such as those of funding bodies or the research team. Even more strongly, non-epistemic interests determine the purposes for which scientific knowledge is used in the CA. What the value-neutrality thesis demands relates exclusively to the CJ: in CJ, science-external values must not play any role. In conclusion, we can formulate and schematically represent the value-neutrality requirement as follows (see [fig. 2.5-1](#)):

(Prop. 2.5-1) *Value-neutrality requirement (V)* : A specific realm of scientific activity, namely its context of justification, should be free from fundamental science-external value assumptions.

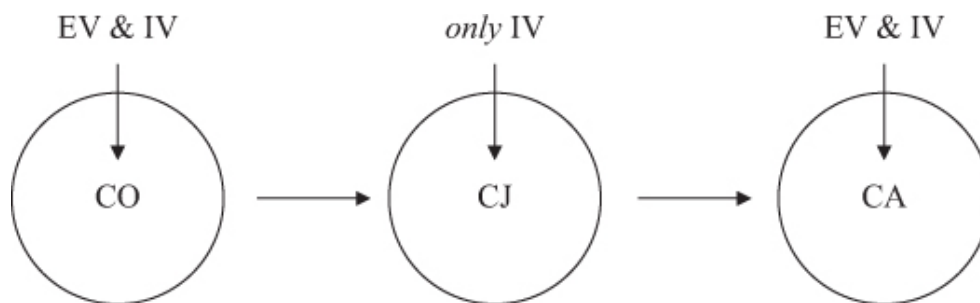


Figure 2.5-1 Schematic representation of the value-neutrality requirement. [↗](#)

Legend: CO – context of origination, CJ – context of justification, CA – context of application; IV – science-internal values, EV – science-external values.

The value-neutrality requirement will hereafter be abbreviated as (V). Note that (V) is a normative requirement, not a factual claim about the actual judgment practices of all or most scientists. There is broad acceptance of the value-neutrality requirement in the sciences, but as noted in [sec. 1.2.9](#) and [2.4.2](#), the requirement is rejected by a significant number of scholars in the humanities and cultural studies (keyword “science wars”). The content of the value-neutrality requirement can also be expressed as a thesis without an explicit normative formulation as follows: the *justificatory status* of scientific claims is *independent* of non-epistemic values. However, even in this formulation, it remains a statement about the epistemic-normative justification status of scientific judgments – and not about the actual judgment practices of scientists.

In our modern, value-pluralistic information societies, politicians and the media rely on scientific legitimations. As a result, scientists face a constant pressure, or at least the temptation, to present recommendations of means in the form of categorical value judgments. Scientists should resist this temptation, as doing so ultimately undermines public trust in science, as recent empirical studies have shown (see [Angeli et al. 2021](#); [Intemann 2024](#), 10). The distinction between factual and value judgments is an essential prerequisite for any credible differentiation between true facts and fake facts, and thus for countering science denial and conspiracy thinking in our time. Only objectivity and value neutrality can enable stable scientific progress through the politically and culturally shifting phases of history.

2.6 The Demarcation Problem and the Methodological Unity of Sciences

2.6.1 Demarcation

If we add the value-neutrality requirement (V) to the epistemic goal (G), the epistemological assumptions (E1–5) and methodological features (M1–4), we can thereby characterize the boundaries of the scientific. As long as legal science, for example, analyzes empirically occurring legal systems, or establishes logical and means-end relationships between norms, it remains a descriptive science. The logical analysis of ethical statements carried out by analytic ethics and meta-ethics also fits easily into the logical-descriptive scientific program. Literary and art theory, which deal with the aesthetic effects of various literary and artistic genres, fit into the framework of empirical-descriptive science, as long as aesthetic effect refers to a *factual* effect on people. Only absolute statements about aesthetic values, which are not relativized to a factually given artistic style, go beyond the scope of descriptive science. This criticism also applies to the tendencies within humanistic psychology and pedagogy which place values like personal development and dignity above objective analysis ([Bühler/Allen 1973](#)). Religious studies, as long as they reconstruct factually occurring religious systems, belong to the descriptive scientific program, while theologies that dogmatically assume the truth of a certain religious, “creed” are departing doubly from the realm of science: first, by assuming the truth of speculations about the nature and existence of God and second by making fundamental value assumptions.

In summary, we can demarcate the boundaries of science as follows: *Sciences in the sense used here are all empirical disciplines in the broad sense (according to G, E1–5, M1–4 and V), together with their formal auxiliary sciences and their means-end-based applied disciplines.*

This characterization of “science” gives us a demarcation between scientific and nonscientific disciplines and fields of mental activity. So, it serves as a suggested solution of the *demarcation problem* (cf. [sec. 1.1](#)). We do not understand this demarcation problem, though, in the same way as logical positivism or critical rationalism did. *First* , we do not identify the nonscientific with the cognitively nonsensical (as e.g. [Carnap 1928b](#)). This identification is unacceptable, because empirically vacuous sentences like “God exists” or fundamental norm sentences like “Help your neighbor” may of course be *understood* , and so have a communicable *meaning* . *Second* , we do not devalue such sentences; we classify them as nonscientific, simply because there is no way to test them empirically. *Third* , we see the significance of the demarcation problem differently from Carnap. Our proposed solution of the demarcation problem does not pursue the inner-philosophical concern of eliminating a part of philosophy called “metaphysics.” Rather, it serves (A) a societal demarcation aim and (B) an evolutionary explanatory concern:

Point A concerns the public debate on the question which kinds of information are sufficiently objective and confirmed to be taught in public schools, and which are not. This requires demarcation criteria between science and religious ideology. As an example, we mentioned in [sec. 1.1](#) two famous cases in which judges based their decisions against the teaching of creationism in public schools on demarcation criteria developed by philosophers of science. However, there is a controversy among philosophers of science themselves as to what the right demarcation criteria should be. Ladyman and Ross ([2007](#) , 30) prefer sociological over logical demarcation criteria. But sociological demarcation criteria are insufficient. For example, Pennock ([2011](#) , 193ff) reports that Judge Jones suggested the application of peer-review methods as a social demarcation criterion. The reaction of Behe, a prominent intelligent design defender, was to promote peer reviews within the creationist camp. Behe pointed to a couple of creationist writers that indeed had been peer-reviewed – not by naturalistic scientists, but by other intellectual creationists. But the consent which a creationist writing earns from other creationists is clearly not an indication of its scientific character.

Point B concerns the *explanation* of the *evolutionary impact* of the sciences. Looking back to the last one or two millennia of human development, science has made extremely rapid progress, compared to ethics, religion, and arts which have developed slowly, at least in those domains where they have not been driven by external technological change. The factual differences in the level of evolutionary impact of the sciences in comparison to ethics, religion or art, and the degree in which they have changed our life, are *so* marked, that they positively beg to be explained. Anyone who wants to understand the powerful independent dynamics of

science and technology has to find out, *what it is* that separates science so strongly from nonscientific systems of ideas. In conclusion: Any philosophy of science which wants to be of practical significance must face up to the demarcation problem.

2.6.2 Methodological Unity instead of Physical Reductionism

Our demarcation account also contributes to the question of the *unity of science*. However, our understanding of the unity of sciences is not ontological, but methodological in nature, and therefore different from traditional accounts. These accounts were predominantly based on *reduction* attempts, that is, on attempts to provide an explanatory or ontological reduction from 'higher-level' or 'special' sciences to *physics*. There are, however, weaker and stronger sorts of reduction accounts. The career of 20th-century reduction accounts started with the program of *physicalism* of Otto Neurath, a prominent logical empiricist, who announced this program as a means to establish the unity of sciences. Neurath's physicalism was weaker than it sounds, insofar, it consisted only in the requirement that all scientific sentences should be translatable into statements about concrete spatiotemporal things and properties ([Neurath 1934](#) ; [Dahms 1994](#) , 170).

A much stronger form of reductionism had been developed in the 1950s and 1960s: the program of *physical micro-reduction*, whose goal was an ontological or explanatory reduction of the laws of special sciences to laws about underlying physical microstructure. ¹

There exist different concepts of reduction. In the standard account of Hempel ([1969](#)) and Nagel ([1970](#) , 125), a physical reduction of a law describing the behavior of a complex system consists in the derivation of this law, or an approximative version of it, from fundamental laws of physics together with contingent boundary conditions, and suitable correspondence principles that translate higher-level concepts into physical concepts. More recent notions of reduction are discussed in van Riel and van Gulick (2025); as these discussions suggest, the core ideas of the Hempel-Nagel concept of reduction can still be considered well-established.

In spite of some positive examples (e.g. the micro-reduction of temperature to mean kinetic energy), this reduction program was not especially successful. It does not describe the way in which special sciences typically proceed. The objects of the special sciences are more-or-less complex systems. At every level of (at least moderately high) complexity, special sciences have discovered their own laws, which are called *system laws* in [sec. 6.5.1](#). Examples are the chemical laws describing the behavior of molecules, biological laws about the functioning of organisms, psychological laws describing human behavior, etc. There exist three arguments, which show that the laws of special sciences are typically *not* reducible to physics, at least not fully (cf. [Cartwright 1983](#) , 104f, 113f):

1. The argument from *hypercomplexity* : The fundamental physical law governing a specific physical system is expressed by its Schroedinger equation. However, already for moderately complex microsystems such as a medium-sized molecule (as opposed to a simple atom), their Schroedinger equation is far too complex to be solvable without radical simplification assumptions. These simplification assumptions do *not* come from fundamental physics but from empirical chemical knowledge. For these reasons, most chemists are convinced that chemistry will never be *fully* reducible to physics (cf. [Hendry 2008](#) , 525ff.; [Mortimer 1986](#)). Much more is this true for the laws that describe highly complex systems such as biological systems: here already the formulation of an approximately complete Schroedinger equation is impossible and attempts to derive biological normal-case laws ([sec. 4.5.2](#)) from fundamental physics and factual descriptions of the physical microstructure of an organism are hopeless. Rather, biological laws are typically discovered by empirical-inductive means. This does not exclude the existence of *partial* reduction relations, for example, between biochemistry and biology or neuroscience.
2. The argument from *multiple realizations* : Empirically, identical macro- properties of a complex system, such as elasticity or intelligence, can be realized by very different physical microsystems ([Fodor 1974](#)). It is hard to find general principles that translate a given macro-property into an exhaustive disjunction of physical micro-properties.
3. The argument from the *systemic properties* : The parts of a complex system are systemic, which means that the properties of these parts in isolation or in a simpler system are different from the properties of the parts within the complex system. An example are the properties of the hydrogen atom H in water

molecules H_2O (O for oxygen), which are different from the properties of the H-atom in isolation, or in the H_2 -molecule. ²

The argument from systemic properties only shows that a so-called horizontal reduction is usually impossible, that is, the explanation of a complex system by the properties of its parts in isolation or in simpler systems. By contrast, “vertical reduction” means that the higher-level properties of a system are reducible to the lower-level properties of its parts *within* the system. Boogerd et al. ([2005](#)) argue in application to biochemical networks that in spite of horizontal irreducibility, vertical reducibility is often possible, and is sufficient for explanations based on biochemical mechanisms. However, since the systemic properties of the parts are determined by the entire structure of the complex system, a micro-reduction of the systemic properties of the parts would lead into an explanatory circle and is therefore not feasible.

These arguments show that the hopes for a physical micro-reduction of the laws of complex systems are dim. For this reason, many philosophers of science have replaced the idea of explanatory reduction by the weaker requirement of *supervenience* ([Hüttemann 2004](#)). Supervenience requires not an explanatory reduction, but merely a “vertical” ontological determination: the physical microstructure of a system *determines* its higher-level properties (at the same given time point). As this determination relation may be infinitely complex, it does not imply the possibility of a reduction by human cognitive resources.

The opposite of supervenience is ontological *dualism*, according to which two physically identical systems may nevertheless differ in some higher-level non-physical properties, for example, “having a mind” according to mind-body dualism. In fact, the supervenience position has been developed by Davidson ([1970](#)) as a naturalistic counter-position to Cartesian mind-body dualism. The supervenience position is fully compatible with the possibility, for which we have argued in this section, that an explanatory reduction of the sciences of complex systems to physics will not be possible in the foreseeable future.

In conclusion, the prospects to establish the unity of sciences by methods of a reduction to fundamental physics are not promising. What one finds along this route is not unity, but *ontological variety*. We suggest that the unity of the sciences should rather be sought by highlighting their epistemic-methodological common ground. The unity of the sciences, if it exists at all, does not exist at the ontological level, but in the form of their *methodological unity*. Of course, there are also important methodological differences between certain disciplines, which will be worked out in several sections of this book ([sec. 1.3](#), 2.4, 3.1, 4.3, 4.5, 5.1, 5.3, 5.5, 6.4, 6.6). But our considerations so far have shown that there exist essential methodological features which are common to all empirical disciplines, despite their differences, and which delineate scientific from nonscientific intellectual activities: the ultimate epistemic goal G, the epistemological assumptions E1–5, the methodological building blocks M1–4, and the value-neutrality requirement V. In this methodological version, the idea of the unity of science can live on in a humbler, but more sustainable version.

2.7 Scientific Inference and Argumentation

In the methodological building blocks M3 and M4, scientific processes of argumentation and inference take place. Thereby, we are dealing with very different processes of inference, which are explained in this section.

2.7.1 Deduction and Induction

When, in accordance with M3, we use a strict law hypothesis together with observation sentences to infer further observation sentences, then we are dealing with a logical-deductive inference. But when, in accordance with M4, we infer from the fact that a large number of actual observation sentences are instances of a law hypothesis, that the law hypothesis is likely to be true, then we have made an inductive inference, or more accurately an inductive generalization. Here are two examples:

(2.7-1)

Deductive inference (logic i.n.s.):

All fish breathe using gills.

Inductive generalization inference:

All fish observed so far (no. 1, 2, ... n)

This animal is a fish.	have breathed using gills.
Therefore, this animal breathes using gills.	Therefore (probably), all fish breathe using gills.
<i>Certain</i> : truth transfer in all possible worlds.	<i>Uncertain</i> : truth transfer only in sufficiently “uniform” possible worlds.

The decisive difference lies in the fact that only deductive inferences transfer the truth from the premises to the conclusion with *certainty*, that is, in all possible “worlds:” *if* the premises (the sentences above the dividing line) are true, then the conclusion (the sentence beneath the dividing line) is also certainly true. Deductive inferences constitute the object of *logic in the narrow sense*. (Some authors place an inductive logic in opposition to deductive logic, for example, [Carnap 1950b](#), sec.43ff; this terminology is disputed, however, see [sec. 5.9.5](#).)

By contrast, inductive generalizations are fundamentally uncertain – the double line in ([2.7-1](#)) indicates this uncertainty. This is because the premises of an inductive generalization tell us only about the cases observed *so far*, while the conclusion generalizes to *all* and in particular to all *future* cases. For this reason, it is also said that inductive inferences are *content expanding*. Only in circumstances or worlds which are sufficiently *uniform*, whose future and past are sufficiently similar, can we reliably infer the truth of an inductive conclusion from the truth of the premises. Nothing can logically guarantee that the future course of events will be sufficiently similar to the course of events observed so far. An extinct fish without gills could, for example, be discovered tomorrow, or a new species of fish without gills could start to evolve through mutation.

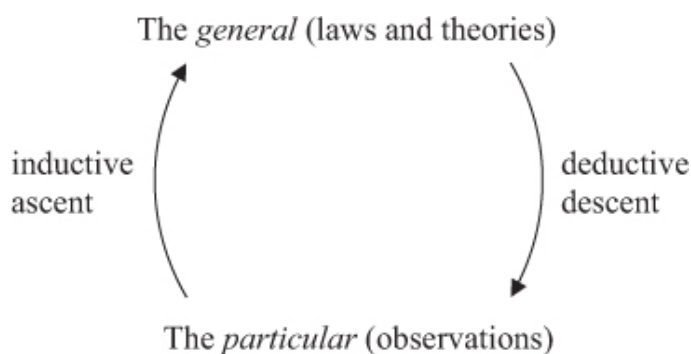


Figure 2.7-1 The inductive-deductive scheme. ↵

This was (very roughly) the main argument of *David Hume*, who was the first philosopher to show that a justification of inductive inference is impossible by means of deductive logic. Hume went on, however, to cast doubt on the rational justifiability of induction as a whole. Hume’s “problem of induction” belongs to the most difficult philosophical problems; a closer look at its subtleties will be given in [sec. 2.11](#).

We have so far explained induction and deduction in the sense of the so-called *inductive-deductive scheme*. According to this scheme, which was already outlined by Aristotle, induction is a method of reasoning (or ascending) from the *particular* to the *general*, and deduction is a method of inferring (or descending) from the general back to the particular (see [fig. 2.7-1](#)).

Although this scheme includes some important methodological features of the sciences, it is *oversimplified* or even incorrect in several respects. For one thing, the binary division of types of inferences on which it rests is *too narrow* – there are far more types of inference, which we will briefly list here. Where the general is concerned, we must distinguish between *strict* (or deterministic) and *statistical* generalizations.

1. Types of deductive inferences :

1.1 From the (strictly) general to the particular – for example, All Fs are Gs, this is an F, therefore: this is a G.

1.2 From the general to the general – for example, All Fs are Gs, All Gs are Hs, therefore: all Fs are Hs.

1.3 *From the particular to the particular* – for example, This is an F or a G, this is not an F, therefore: this is a G.

By contrast, there are no deductive inferences from the particular to the (essentially) general.

2. *Types of inductive inferences* (see also [Carnap 1950b](#), 207f):

2.1 *From the particular to the general* :

2.1.1 *The strict inductive generalization* : All Fs observed so far were G, therefore (probably), all Fs are Gs.

2.1.2 *The inductive-statistical generalization* : r % of all Fs observed so far were G, therefore (probably), approximately r % of all Fs are Gs.

2.2 *From the particular to the particular – the inductive prediction* : All Fs observed so far were G, therefore, the *next* F will probably be a G, too.

2.3 *From the general to the particular – the inductive-statistical specialization* : r % of all Fs are Gs, this individual is an F, therefore, we can be r % sure that this individual is a G.

Inductive specializations are also called “direct inference” ([Levi 1977](#)) or “statistical principal principle” ([sec. 3.13.4](#)). Two facts about them are remarkable: First, the strict variant of inductive specializations is deductively valid. Second, like inductive generalizations, inductive-statistical specializations assume uniformity: they are only reliable if our actual observation sample is representative for the total domain.

2.7.2 *Popper and the Significance of Inductive Inference in the Sciences*

The founder of critical rationalism, Karl Popper, advocated the thesis that the induction problem is, on the one hand, unsolvable, but, on the other hand, that empirical science can proceed without inductive inferences altogether (1935/2002, sec. I). In the following, we want to demonstrate, based on a differentiation between *three kinds of induction*, that Popper’s arguments bring to light important insights, but that his radical claim that empirical science could go without induction is not tenable (cf. [Schurz 2002a](#)).

1 *Methodological induction*. This view of induction connects directly with the inductive-deductive scheme explained in [sec. 2.7.1](#) . According to this view, induction is a method of inferring general laws and theories by inductive generalization from particular observations. Popper criticizes the view contained in methodological induction that induction is the central method of discovering or generating law hypotheses and theories. The belief that such a discovery method is at all necessary in science rests, says Popper, on a confusion of the *context of discovery* and the *context of justification* (1935/2002, sec. I.1–2). How scientific hypotheses are generated, be it through induction, intuition, or trial and error, is completely *irrelevant* to the context of justification. Seen like this, methodological induction is *dispensable* in the context of the justification of science.

By contrast, call to mind [sec. 2.1](#) : not all but *some* methods of generating hypotheses may give us a certain measure of justification as well. This is the case, for example, for the inductive generalization inferences explained in [sec. 2.7.1](#) : these inductive inferences are methods of discovering empirical law hypotheses, which at the same time inductively justify these law hypotheses, that is, preliminarily confirm them. More refined inductive discovery algorithms can be found in Langley et al. ([1987](#)) (see also [sec. 4.2.2](#) , 4.3.2). Hypothesis generating procedures which at the same time give us a justification are of course cognitively much more *efficient* than blind trial and error procedures.

In two important points, however, Popper’s criticism of methodological induction is right. (i) Hypotheses do not *have* to be produced by methodological induction – all methods of producing hypotheses are allowed and deserve critical assessment. (ii) Inductive generalization methods work for empirical law hypotheses, but they are no longer sufficient for the construction of scientific theories. Even very simple theories – like, for example, the Galilean law of free fall, which posits the idealization of a friction-free fall – cannot be created by inductive generalization procedures alone, since they contain concepts that transcend the observable – so-called *theoretical concepts*, such as “friction-free movement,” “mass,” or “force.” So Popper’s criticism of

the possibility to create scientific theories by methodological induction (see 1983, 118) is to a large extent legitimate.

2 Logical induction . This view of induction goes back to Carnap ([1950b](#)) and understands induction not as a method of discovery, but as one of justification of scientific hypotheses, that is, as a method of determining the *degree of confirmation* conferred on theories by given observational data. This should happen by the establishment of a system of logical probability, in which the inductive probability of a theory given a body of observational evidence may be determined in a rational way (see [sec. 4.7](#)). Instead of the inductive degree of confirmation, one can alternatively regard the degree of (epistemic) *truthlikeness* of a theory relative to a body of evidence as the central measure of evaluating theories (see [sec. 5.11.1](#)).

Popper's criticism of the view of logical induction now says the following (1983, 19f, 23): while we are able to determine that relative to given observations, a theory T_1 is *better* confirmed or corroborated than another theory T_2 , or that T_1 is closer to the truth than T_2 , we still cannot give an *absolute* value for the probability or truthlikeness of T_1 or T_2 , as tomorrow a new superior theory T_3 may be developed, which completely overshadows both T_1 and T_2 . Popper's argument represents a substantial insight into the *limitations* of logical induction. In order to assign a nonarbitrary *numeric value* to the inductive probability or truthlikeness of a theory T given some evidence, one must have at one's disposal, among other things, a comprehensive list of all relevant *alternative theories* $\{T_1, T_2, \dots\}$. However, to establish such a list in a nonarbitrary manner is, at least in the case of theories, impossible. For any given theory T , there are always indefinitely many alternative theories, as these may contain any number of new theoretical concepts. A measurable linguistic space containing all possible alternative theories simply does not exist (see [Howson 2000](#) , 46). Many theories that were later accepted were still unimaginable a century earlier in the history of science (see [Stanford 2006](#)).

Again, in the case of *empirical* hypotheses, the situation is different. Here, it is possible to establish a finite list of all possible alternative hypotheses in a *limited* observational vocabulary. Let us consider, for example, the set of all interval hypotheses of the form $H_n =$ "between n and $n+1\%$ of all people are brown-eyed," with $0 \leq n \leq 99$, and let us assume a uniform distribution of the initial probabilities of these hypotheses. Then, the inductive probability of H_n , given a sample of say m people, including $k \leq m$ with brown eyes, can be calculated based on Kolmogorov's probability axioms and the statistical principal principle ([sec. 4.7.2](#) – [4.7.3](#)). In summary, Popper's criticism of logical induction for empirical law hypotheses is unjustified or at least exaggerated; but it is certainly justified in the case of scientific theories.

Based on Popper's criticism of logical induction, we can evaluate competing theories only *comparatively* , that is, in relation to each other. Once, we have carried out an empirical success comparison of currently available alternative theories, we will base our future predictions and actions on those theories which up to now have been *most successful* , in other words, on the ones which have been corroborated best. Popper, of course, also sees it this way (1983, 65; 1979, sec. I.9). It follows from this that Popper's "deductivist" program of corroboration, too, contains a fundamental inductive step – a third kind of induction, that has been called epistemic induction ([Musgrave 2002](#)) or meta-induction ([Schurz 2008b](#)):

3 Epistemic induction or meta-induction says the following: If a theory T_1 has been more successful than T_2 *so far* (explanatory and prognostically), then it is probable, *relative* to the given state of evidence, that T_1 will also be more successful than T_2 *in future* . In other words, the success preferences established *so far* are projected inductively into the future. This induction principle is called "epistemic" or "meta-," because it is not object-level hypotheses (about ordinary events), but epistemic meta-hypotheses about the degree of corroboration of our object-hypotheses, which are being inductively projected. The epistemic induction principle is indeed indispensable for all empirical disciplines. Were this principle not to be accepted, there wouldn't be any point in Popper's method of testing. Success so far would then simply be irrelevant to our future actions. Although for example the view that heavy bodies on Earth are disposed to fall to the ground and not to float freely has been better corroborated so far than the opposite view, this would not be a reason to make this view the basis of our future actions, since logically, all bodies might just as well start floating as of tomorrow. Unfortunately, Popper kept an exaggerated distance to logical empiricism throughout his life,

and condemned any form of induction whatsoever. But his arguments on this topic in (1983, e.g. 66f) seem to rest more on mere terminological maneuvers than on actual material reasons (see [Schurz 2002a](#)). Other critical rationalists like Watkins ([1984](#) , 340f) and Musgrave ([2002](#)) have accepted the epistemic induction principle.

Epistemic induction, let us remember, is always only *comparative* . An important consequence of this is the following: according to epistemic induction, the degree of confirmation of a theory is always *doubly* relative:

- relative to the given state of observational knowledge, and
- relative to the current state of alternative theories.

2.7.3 Abduction and Inference to the Best Explanation

Scientific theories differ from empirical law hypotheses in that they contain theoretical concepts which speak about unobservable entities or properties. Inductive generalization enables us to infer from particular observations to empirical law hypotheses, but not from empirical law hypotheses to scientific theories. This is because inductive inferences cannot introduce new (theoretical) concepts into the conclusion, which are not contained in the premises. There is, however, one further – but controversial – kind of inference, by which this is possible, namely abduction, or inference to the best explanation. Roughly expressed, this is an inference from an observed *effect* to a likely *cause* , for example, from a curved trail in the sand to the recent passage of a sand viper.

At this point, we must warn against a *terminological misunderstanding* . A number of authors understand the concept of induction in a *broad* sense – they identify inductive inferences with *any* kind of content-expanding or nondeductive inferences, and so regard abduction as a form of induction (e.g. [Pollock 1986](#) , 42; [Bird 1998](#) , 13; [Ladyman 1998](#) , 28). By contrast, for the sake of a clear demarcation between distinct concepts, we understand induction always in the above outlined *Humean* sense, and see abduction as a second nondeductive or content-expanding kind of inference.

The abductive form of inference goes back to C.S. Peirce. In (1878a), Peirce explicated abductive inference formally as a retrodictive inference from “all Fs are Gs” and “this is a G” to “this is an F.” While this characterization of abduction does dominate the computer science literature on abduction (see [Flach/Kakas 2000](#)), it is much too narrow to grasp what the late Peirce was trying to achieve. The essential thing for Peirce ([1903](#) , §170) was that abductive inference allows new theoretical concepts and models to be introduced in the sciences. Newton, for example, inferred abductively from the movement of the planets around the Sun to the existence of a gravitational force. The epistemological inference from subject-internal perceptual experience to external empirical facts is also an abductive inference: the assumption that my visual experiences are being caused by objects existing independently of me is the best explanation for the constantly recurring regularities in my visual experiences (cf. [Schurz 2024 a](#), ch. 9). As is shown in Schurz ([2008a](#)), abduction comprises a whole *family* of kinds of inferences, which have the following pattern in common (cf. [Peirce 1903](#) , §189; [Niiniluoto 1999](#) , S439):

(Prop. 2.7-1) *General pattern of abduction* :

Premise 1 : A (singular or general) fact E, in need of explanation.

“ *Premise* ” 2 : A background knowledge K, which implies for a hypothesis H, that H is a potential and sufficiently plausible explanation for E.

Abductive conjecture : H is true.

As Peirce himself emphasized, the trustworthiness of an abductively acquired hypothesis is rather preliminary: the abducted hypothesis must go on to be empirically tested by deduction and induction in order to acquire the character of a probable hypothesis (1903, §171). More importantly, the given background knowledge always contains *several* possible hypotheses, H_1 , ..., H_n , which *potentially* explain the given evidence E (i.e. if they were true, they would explain E); and the abductive inference selects the most

plausible hypothesis among them. In this sense, Harman ([1965](#)) transformed Peirce's concept of abduction into the following inference to the best or best available explanation (IBE):

(Prop. 2.7-2) *Inference to the best (available) explanation – IBE* : Infer abductively to the most plausible explanatory hypothesis among all explanatory hypotheses for the given explanandum E that are possible, or are available in the given background knowledge, respectively.

Since the later Peirce also regarded abduction as IBE (1931ff, CP 6.525), abduction and IBE can be seen as approximately equivalent forms of inference (cf. [Douven 2011](#) ; [Williamson 2016](#) ; [Niiniluoto 2018](#) , 17f); however, this equivalence thesis is controversial (cf. the discussion in [Schurz 2023c](#) , ch. 1).

According to the understanding of Harman's ([1965](#)) and other representatives of IBE ([Lipton 1991](#) ; [Armstrong 1983](#)), an IBE infers from the explanatory success of a theory to its *likely truth* in the realist sense. Lipton ([1991](#) , 58) and van Fraassen (1989, 142ff) have pointed out that one can infer only to the best *available* explanation, because no one knows all possible explanations. In this situation, Kuipers ([2000](#) , sec. 7.5.3) has made an interesting suggestion. He brings the same objection into play as Popper made against logical induction: De facto, we can only ever reach *comparative* judgments, according to which a theory is explanatorily more successful and, hence, likely to be closer to the truth than all alternative theories *known* to us. Since this explanatory success includes also predictive success, Kuipers speaks more generally of "empirical success." Kuipers therefore suggests modifying the pattern of abduction in the following way: first, instead of speaking of an "inference to the best explanation," he speaks of an "inference to the best theory," and second, that a theory T is empirically most adequate among all known alternative theories does not lead to the conclusion that T is likely to be true, but only that T is *closest to the truth* among all known alternative theories. The interplay of the epistemic inductive inference and the abductive inference to the best theory is shown in [fig. 2.7-2](#) .

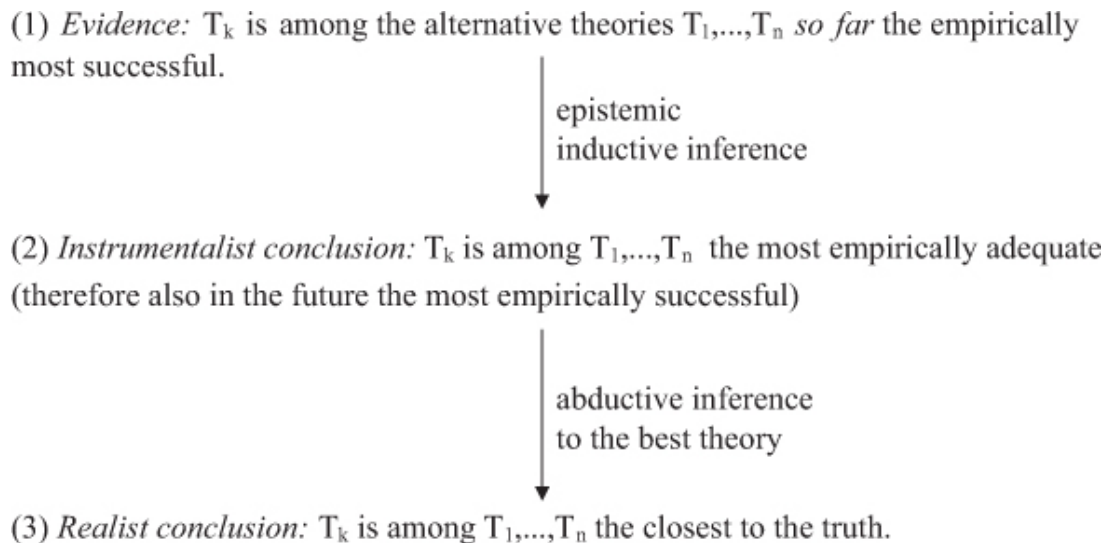


Figure 2.7-2 Interaction of epistemic induction and abduction. ↵

In [fig. 2.7-2](#) it becomes clear how these two kinds of inference are connected with the positions of *instrumentalism* and *realism* in the philosophy of science (see [sec. 5.7](#)). Instrumentalists, like for example van Fraassen, accept the inference from (1) to (2), but reject the abductive inference from (2) to (3), because according to them, theories may only be justified to be more or less empirically adequate, but not to be true or false in the realist sense (van Fraassen 1980, 11f, 20; 1989, 142f). Realists in the philosophy of science believe on the other hand in the likely truth or truthlikeness of well-confirmed theories; for them, it is that very inference from (2) to (3) which is essential.

2.7.4 Monotonic and Non-monotonic Inferences

Because deductive inferences are certain, that is, preserve truth (from premises to conclusion) in all possible worlds, they are monotonic in the following sense:

(Def. 2.7-1) *Monotonicity* : A valid inference (from premises $P_1, \dots, P_n$ to a conclusion C) is called monotonic, iff it remains valid after the addition of arbitrary further premises.

By contrast, inductive or abductive inferences don't possess the monotonicity property, because they are uncertain, that is, they don't transfer truth in all, but only in sufficiently "normal" worlds. We restrict the notion of *validity* to deductive inferences (indicated by a single forward slash "/"), while in the case of uncertain inferences, we speak instead of their (nondeductive) *correctness* (indicated by a double forward slash "//"). With this notation, monotonicity means that if " $P_1, \dots, P_n / C$ " is valid, then " $P_1, \dots, P_n, Q / C$ " is valid for every new premise Q .

The non-monotonicity of uncertain inferences makes it possible that " $P_1, \dots, P_n // C$ " is correct, but " $Q, P_1, \dots, P_n // C$ " is incorrect for some (true) premise Q , where Q typically expresses "exceptional information." For example, "1000 swans have been observed to be white // Therefore probably: All swans are white" is a correct inductive generalization, but the premise-expanded inference "1000 swans have been observed to be white and 10 to be black // Therefore probably: All swans are white" is inductively incorrect. The logical examination of *non-monotonic* inferences has become a significant area of inquiry since the 1980s (cf. [Gabbay et al. 1994](#) ; [Schurz 2004](#)).

A non-monotonic inference with true premises has only *preliminary* value and can be made worthless by new (true) information. In epistemology, non-monotonic inferences are therefore also referred to as *defeasible* (cf. [Pollock 1974](#) , 36ff). The preliminary nature of non-monotonic inferences is expressed by the fact that they have a different *epistemic application rule* than monotonic inferences. Here, we take the *belief system* (or 'knowledge' system) S of a person or group of persons to be the set of propositions which are justifiably accepted as true by these persons.

(Prop. 2.7-3) *Application rule for monotonic inferences* : If a deductive inference is valid, then you should add its conclusion to your belief system S , as long as S contains the premises.

(Prop. 2.7-4) *Application rule for non-monotonic inferences* : If an uncertain inference is correct, then you should add its conclusion to your belief system S , as long as S contains the premises, and these premises cover *all* propositions in S which are of relevance to the conclusion. For inductive-probabilistic inferences, [prop. 2.7-4](#) is equivalent to Reichenbach's rule of the narrowest reference class and Carnap's requirement of total evidence (see [sec. 3.9](#) , [3.13.4](#) , [6.3.2](#)).

Complementary and Advanced Topics

2.8 How Much Metaphysics is Presupposed by Science?

2.8.1 Truth, Idealism, and Realism

One may object that our reconstruction of the supreme goal (G) of the sciences and the assumption ($E1$) of minimal realism assumes too much of realism. Famous scientists and philosophers of science, such as George Berkeley (1685–1753), Ernst Mach (1836–1916), and the early Rudolf Carnap, were idealists or positivists, that is, anti-realists. In their view, science is possible if it is merely based on human intersubjective experiences, *without* the presupposition of a subject-independent reality that transcends these experiences. We do not deny this possibility.

Admittedly, our assumption ($E1$) makes a (weak) realistic assumption, which is discussed below. Some philosophers argue that already our explication of the goal (G) of science, the discovery of content-rich

truths, presupposes realism, insofar as truth means correspondence of statements with an assumed reality. But this is not the case, since (G) does not say that the domain of phenomena to which true statements correspond has to belong to an external and subject-independent reality, as claimed by realists. The domain of phenomena called ‘reality’ may also consist of experiences, as assumed by idealists.

The major epistemological positions can be summarized as follows (see [Dancy 1985](#) , 136ff; cf. [Schurz 2024a](#) , [sec. 2.7](#)). The domain of entities to which true statements correspond consists either of

- a. one’s own subjective experiences (*subjective idealism* or solipsism),
- b. the experiences of all subjects (*intersubjective idealism*),
- c. the *possible* experiences of all subjects (*possibilistic idealism*), or
- d. a subject-independent reality that causes the experiences (*realism*).

Solipsism is a rather extreme position (most idealists have been possibilistic idealists). Be that as it may, the important point to observe here is that the concept of *truth* makes sense in *all* of these positions. Hence, truth in the sense of correspondence theory is not realistically biased but *metaphysically neutral* (cf. [Kirkham 1992](#) , [sec. 6.5](#)). So also, our formulation of the supreme goal (G) of science is neutral. It does not entail that successful science presupposes the epistemology of realism. The goal of truth makes sense for *all* of the epistemological positions; only the *kinds* of truths that, according to these positions, are pursued by science are different (true descriptions of experiences versus true descriptions of subject-independent facts).

While (G) is metaphysically neutral, our epistemic assumption (E1) is not. It asserts that realism, in the explained weak version of it, is the *normal* epistemological assumption of science. I think this claim is defensible: more or less all empirical scientists assume the external reality of their objects of investigation; idealist or anti-realist worldviews are the clear exception in normal science, although the realistic attitude of normal science can be rattled in scientific revolutions (recall [sec. 2.1](#)).

What distinguishes our minimal realism from all sorts of *metaphysical realism* is that neither the existence of a subject-independent reality, nor its knowability, is assumed to hold on a priori reasons or is a necessary presupposition of knowledge. By contrast, the minimal realism of assumption E1 is a *hypothetical-constructive* realism, insofar as we consider realism not as a presupposition, but as a hypothesis or model construction – in analogy to scientific hypotheses. Our realism is committed to *methodological naturalism* , as the hypothesis of realism must be corroborated by the *empirical success* of our cognitive belief system, as the *best explanation* of this success. This abductive form of realism corresponds at the same time to a variant of empiricism, namely *abductive empiricism* , which – unlike the reductionist empiricism of the neopositivists – is compatible with realism ([Schurz 2024a](#) , 90).

Because of its naturalistic tenets, hypothetical-constructive realism is compatible with legitimate doubts about a ‘full’ realism found in the philosophical interpretations of *quantum mechanics* . In all of its interpretations, quantum mechanics assumes the existence of a microphysical reality that possesses some subject-independent and discoverable properties (so-called observables). Only certain properties of entangled (or superposition) states are not objectively measurable, because it is a postulate of quantum mechanics that every measurement transforms an entangled state into a statistical ensemble of pure states (see [Herbert 1985](#) , sec. 8+9; [Penrose 2007](#) , sec. 29).

Hypothetical-constructive realism is also compatible with *structural* realism. In its *epistemic* version, structural realism says that only the structure but not the intrinsic nature of the entities making up reality are knowable ([Worrall 1989](#) /1997; [Psillos 1995](#) ; [Votsis 2011](#)). The *ontic* version of structural realism is motivated by quantum mechanics and claims that even the entities that make up reality are themselves not individuals but structures ([Ladyman 1998](#) ; [French/Ladyman 2003](#)).

2.8.2 Constructive Realism versus Radical Constructivism

Hypothetical-constructive realism differs markedly from radical *constructivism* mentioned in [sec. 1.3.9](#) . The latter position was developed by psychologists such as Glasersfeld ([1995](#)) or biologists like Maturana and Varela ([1992](#)). Historically speaking, radical constructivism is a modern variant of subjective idealism. The main argumentation of this philosophical position is based on the fact that our perceptions and concepts of reality are cognitive *constructions* of our mind or brain, respectively. Radical constructivists, however, infer from this hardly deniable premise that even ‘reality’ itself is a construction, that is, a reality that is

independent of our constructions does not exist. A similar argument has been put forward by social constructivists such as Latour and Woolgar ([1986](#)), in regard to the ‘construction’ of scientific facts by scientists.

In this subsection, we want to show that the argument of radical constructivists is based on a *fallacy*. We reconstruct this argument as follows (e.g. [Glaserfeld 1995](#), 40, 51; [Kriz et al. 1990](#), 19f):

(2.8-1) *Premise of radical constructivism = epistemic constructivism* : Our perception and mental representation of reality is not something given, but the result of an active cognitive construction.

Conclusion of radical constructivism = ontological constructivism (idealism) : Therefore reality, too, is not given “in itself,” but “constructed” or “constituted” by us.

The crucial fallacy of this argument is the ‘magical transformation’ of epistemic into ontological constructivism (cf. [Sokal/Bricmont 1998](#), 102). Hypothetical-constructive realism shares the premise of ([2.8-1](#)) with radical constructivism, but it definitively rejects the conclusion. The *invalid* step from epistemic to ontological constructivism is frequently substantiated by constructivists with the help of the following ‘justification’ (cf. [Kriz et al. 1990](#), 20; [Glaserfeld 1995](#), 51), which may be called the *Kantian fallacy*, because it corresponds to a popular interpretation of Kant’s philosophy:

(2.8-2) “ *Justification* ” of radical constructivism : It is impossible to know something about reality *in itself* – independent of its epistemological construction. The only thing about which we can know something is the reality *constructed by us*.

The subtle error of this justification attempt is that the concept of “knowing” is understood in the *naïve-realist* sense of a kind of direct mirror-imaging. Then it is indeed impossible to know something about an object that is “independent of our epistemological construction” – which in this case means: “ *not influenced* by this construction.” Modern truth theory, though, understands the relationship between a true statement and its object not as a quasi-identical mirror-image, but as a *structural correspondence*, which *conveys certain information* and has to be neither complete nor unique. In this view, there is no incoherence in saying that *by means of* our epistemological constructions we know something about a reality which exists *independently* of our epistemological constructions. For the independently existing reality is not taken to be immediately given to our mind, as in metaphysical or ‘direct’ realism, but is postulated as a system of hypothetical entities, whose existence can best explain the factual success of our knowledge. Giere ([1999](#), 174ff) and Kuipers ([2000](#), sec. 13) have called this position “constructive realism,” and we call it *hypothetico-constructive* realism or alternatively *abductive* realism, in order to emphasize that realism too is ultimately a fallible hypothesis.

In conclusion, it is the very argumentation of radical constructivists which clings to naïve-realism, inasmuch as it uses a naïve-realist interpretation of the objects of our knowledge to prove that these objects are subjectively constructed; while from the hypothetical-realist view, the statement ([2.8-2](#)) and so the whole argumentation of radical constructivists is untenable.

2.9 On the Theory (In)dependence of Observation

Classical empiricists argued that all scientific (or cognitively valuable) concepts are reducible to observational concepts. That this is not so was convincingly demonstrated by the recognition of *theoretical concepts* in the middle and later stages of logical empiricism (cf. [Carnap 1936](#) /37, 1956a). Theoretical concepts go beyond what is observable – this is a generally accepted insight in post-positivistic philosophy of science ([sec. 5.1](#)). Accordingly, the empiricism in our reconstruction of the common epistemological ground of science ([sec. 2.2](#)) is reduced to a minimalistic version according to which theories, although going beyond what is observable, must at least have *some* observable consequences by which they are testable.

What minimalistic empiricism still assumes is the existence of at least *some* theory-independent observation concepts and sentences, which can be used to independently test competing theories. However, even this minimalistic empiricism came under attack in the post-positivistic philosophy of science. The thesis of the *theory-dependence* of all observation and experience has been made popular by Hanson ([1958](#)), Kuhn ([1962](#)), and Feyerabend ([1975](#)); with precursors in Wittgenstein ([1945](#) /2001, part 2, §xi) and Popper (

[1935](#) /2002, new appendix X). Meanwhile, this thesis seems to be accepted by the majority of philosophers of science. Closer inspection reveals, however, that what is accepted by the majority of philosophers is a cautious and loose version of this thesis, that leaves plenty of room for interpretations that are compatible with our epistemological assumptions (E3) of objectivity and intersubjectivity, and (E4) of minimal empiricism. For many philosophers of science, what is important in the theory-ladenness thesis is a critical attitude toward what is called “data” in science: these data are typically theory-laden interpretations of the scientist’s perceptions, and the assumptions made in these interpretations are often concealed behind a veil of expertocracy. Nevertheless, I will try to show in this section that in all empirical sciences, one can find theory-neutral observation concepts and statements that scientists may draw on in cases of conflicts of opinion.

Before I start my enterprise, let me try to demonstrate the significance of this question for the philosophy of science. A *complete* theory-dependence of observations, in the sense that *all* observations are theory-dependent, would turn empirical science into a *circular* affair without any hope of objectivity (cf. [Nagel 1979](#) , 79). For suppose that two adherents of rival theories T_1 and T_2 are observing the same area of space-time, and assume that their respective theories imply the contradictory predictions P_1 and P_2 . If complete theory-dependence were given, then the observations of the theory-adherents would be determined by their theoretical *expectations*. The supporter of T_1 would then observe P_1 where the supporter of T_2 observes P_2 . So, each prediction would be a *self-fulfilling* prediction, and an independent empirical test of theories would be impossible.

Complete justificatory circles are demonstrably epistemically worthless (see [sec. 2.11](#)). By contrast, a *partial* theory-dependence, in the sense that there are at least some theory-neutral observations on which theory-dependent scientific data rest, is not a threat to the possibility of objective empirical tests. Adam ([2004](#)) and Brown ([1993](#)) have studied cases in which the recognition of a scientific datum D that falsifies a theory T relies on the assumption that theory T is true. This sounds paradoxical, but in fact is not: the case is harmless, because the datum D was gained from a theory-neutral observation O using T . This means (in the simplest case) that T entails the implication *if O , then D* . Moreover, that T is falsified by D means that D entails not- T . Taken together, this implies by simple logical means that O entails not- T , that is, that the observation O falsifies T . Nothing paradoxical is involved in this situation.

In conclusion, for a defense of the objectivity of empirical science, it is sufficient to refute the thesis of complete theory-dependence, that is, to establish the existence of *some* theory-neutral observation concepts and statements in science. In the next subsection, I will sketch eight central arguments in favor of different forms of theory-dependence as well as the corresponding counterarguments. I will try to show that although the pro arguments make correct points, they do not exclude the existence of theory-neutral observations in the narrow sense of perceptions.

2.9.1 Theory-Dependency of Observation – Pros and Cons

Pro-1 : Observations are expressed as realist statements . Observation statements such as “this thing is green” are usually expressed in the realistic mode. So, they imply the *hypothesis of realism* , that is, the existence of subject-independently existing objects (“this thing”) that cause our subjective perception (cf. [Popper 1935](#) /2002, sec. V, §25, for a version of this argument). Observation sentences in this realist interpretation transcend the content of subjective experience. We can be in *error* about them, for example, because our sense experience may rest on a *hallucination* .

Contra-1 : This argument is true, but this is not what is meant with theory-dependency of observation. What is meant is that the *content* of one’s observations depends on the *content* of one’s background theory (expectation or worldview). But the question whether one is an epistemological realist or a phenomenalist idealist does not affect the content of one’s observations. It merely changes the interpretation of the term “thing:” in the realist interpretation, this term refers to assumed entities outside of one’s experiences, and in the phenomenalist interpretation, it refers to entities within one’s experiences. But whether one is a realist or an idealist does not turn a green-observation into a blue-observation (this argument can be found in [Carnap 1928b](#)).

Pro-2 : Observations are theory-guided . The selection of relevant problems and experiments is essential in every scientific research, for reasons of complexity alone. This selection is guided (among other things) by one's theory: which theory I support determines which observations I look for. ³

Contra-2 : It is true that the selection of one's observations depends on one's theories, and more generally, on one's interests. But that observations are selected in a theory-dependent way does not imply that the observations *themselves* are theory-dependent, in the sense that supporters of different theories would observe different things when perceiving the same spatiotemporal region or measurement result. Supporters of the theory T_2 *could* also make the observations made by supporters of the theory T_1 and *would* then recognize their truth. More generally speaking, the fact that empirical research is interest-driven does not imply that certain observations *cannot* be made, but at most that these observations are not *wanted* to be made, or that they are *ignored* . However, such a dogmatism is a sign of *bad* science and violates the principle of *total evidence* ([sec. 3.13.4](#)).

Pro-3 : Perceptions are unconscious theoretical constructions. That perceptions or observations in the *narrow* sense are unconscious theory-like constructions is an insight that goes back to Helmholtz ([1896/1910](#)). The brain's visual regions use refined neural construction mechanisms to create a consciously experienced 3D image from the unconscious 2D retinal light-impressions. Normally, this construction process generates a unique 3D image within milliseconds, so that its constructive character is *hidden* from the consciousness: you think you see "reality itself" (for this reason, humans are intuitively naïve-realists). But in certain – typically in artificially engineered – situations, this visual construction process reveals its constructive character, because it leads to consciously experienced visual *ambiguities* , *illusions* , or other strange effects. This has been impressively demonstrated by the findings of cognitive psychology (cf. [Rock 1984](#)). Simple examples of ambiguities are *tilting pictures* such as Jastrow's duck-rabbit, a picture that can be seen either as a duck or a rabbit (see [fig. 2.9-1](#)). The duck-rabbit has been used by Wittgenstein ([1945/2001](#), part II, §xi), Hanson ([1958](#) , 12), and Kuhn ([1962](#) , 112ff) as evidence for the 'theory-dependence' of perception, insofar as what we *expect* to see influences what we see.



Figure 2.9-1 Duck-rabbit. [↗](#)



Figure 2.9-2 Mueller-Lyer illusion [↗](#)

An example of an *optical illusion* is the Mueller-Lyer illusion in [fig. 2.9-2](#) . The explanation of this illusion lies in mechanisms of 3D-vision: the arrows and inverted arrows produce the 3D-effect that the right stroke appears to be closer to the observer than the left stroke and is therefore perceived as being shorter, although both strokes have equal length (see [Rock 1984](#) , 163, 167).

Contra-3 : Neither the logical empiricists nor Popper had good replies to the psychological arguments concerning the constructive character of perception, because according to their position of "logical purity," philosophy should be independent from psychology. What counts as an observation sentence was logically speaking a matter of "pure convention" (cf. [Popper 1935/2002](#), sec. 5, §30; [Neurath 1934](#)). We now present some arguments against the theory-ladenness of perceptions that are themselves based on findings from cognitive psychology.

The findings about the constructive character of visual processes refute only *direct realism* ([Dancy 1985](#) , 144). According to this epistemological position, we see things directly, as they are 'in themselves,' without

any mediating processes that construct intermediate images. In the light of contemporary knowledge about the neuro-cognitive process of vision, direct realism seems to be an untenable position (see [Schurz 2024a](#) , [sec. 3.5](#)).

However, all findings of cognitive psychology also confirm that our visual perception processes and their results are quite stubbornly *independent* of our *learned* information. Everyone experiences the right stroke in [fig. 2.9-2](#) as shorter than the left one, even if she has been informed that this visual experience is an illusion. These facts motivated Fodor's thesis of the modularity and impenetrability of vision ([Fodor 1984](#) , 34ff). Also, the Helmholtzian constructivist Rock ([1984](#) , 228) acknowledges the independence of vision from acquired information.

In the tilting picture of [fig. 2.9-1](#) , everybody is able to perceive it either as a duck or as rabbit and to switch back and forth between these two gestalts. Kuhn argued that a similar gestalt switch took place when astronomers saw Uranus before 1781 as a fixed star, but after 1781 as a moving planet (1962, 111, 121f). But it does not seem right to conceive the Uranus case as perceptual switch: The visual content of the Uranus-perception is always a bright spot in the sky, which is perceptively stable and does not switch. The question whether it represents a fixed star or a planet is not a question of perception but of theoretical interpretation.

The processes of human vision are based on innate mechanisms that have developed in millions of years of evolutionary history. Their selection history explains why the *same* mechanisms, which under normal environmental conditions give a correct visual depiction of reality, can create optical illusions under certain non-normal conditions. Evolutionary considerations also explain why human perception, in order to be evolutionary successful, must be independent from expectations, for otherwise it could not correct *wrong* expectations. As an example, consider prehistoric men entering a cave which is inhabited by a cave bear that they did not expect: if their perception would not immediately have corrected their wrong expectation, these prehistoric men would not have survived.

By "theory-dependence" of perception, one could at best understand that perception is dependent not on learned information, but on 'innate theories' which all humans share, because they are rigidly wired up in their visual cortex. However, we do not want to understand the concept of "theory" in such an overstretched way. By the theory-dependence of perception, we always only mean the dependence of perception on *acquired* information that may vary in different subjects, cultures, or scientific communities. It seems that visual perception in this sense is – at least to a large extent – theory-independent.

Pro-4 : Sometimes, visual perception depends on one's pre-expectations . The last statement of Contra-3, that visual perception is independent from pre-expectation, is not generally true – it holds only for clearly presented visual stimuli. There exist some psychologically well-investigated experiments, to be discussed below, in which visual perception does indeed depend on pre-expectations.

Contra-4 : This is only the case for the stimuli at the perceptibility threshold , that is, when the limits of perceptual recognition capacities are reached (cf. [Brewer/Lambert 2001](#) , S181). These limits are reached in three kinds of cases:

1. *Poorness* of the stimulus: For example, (1.1) in a dark forest at night, anxious persons seem to see the contours of mysterious figures in the dark thicket, which disappear at daylight. Or (1.2), if the relevant stimulus is too small to be seen, for example, when you ask which of two approximately equally long strokes is the longer one, the test persons' answers may become expectation-dependent. The perceptibility threshold is also reached when (1.3) the presentation time of the stimulus is very short, as in high-speed reading tasks (cf. [Estany 2001](#) , 208).
2. *Ambiguity* of a stimulus that has no unique gestalt solution, such as the duck-rabbit in [fig. 2.9-1](#) : Here, one's expectation about the gestalt influences which gestalt one sees first (though one may switch the gestalt at will).
3. A *high complexity* of the relevant stimuli patterns is a third reason, not for expectation-dependence, but for dependence on acquired knowledge: here, visual *training* may be necessary to detect the relevant visual patterns. As an example, think of the radiologist who recognizes pneumonia in the X-ray image of your chest.

Because of these finding, some cognitive psychologists described human vision as a combination of bottom-up processes (information flow from the retinal cells) and top-down processes (information flow from higher

cognitive regions; cf. [Estany 2001](#) , sec. 4). Although this description is not false, it misses the important point of this combination, namely the *priority ordering* in favor of bottom-up information. Top-down information from expectation by memory can correct visual bottom-up information *only* if this bottom-up information is *too poor* to allow a clear gestalt interpretation. Otherwise, bottom-up information *overrides* top-down expectations. This is illustrated, for example, by an experiment in which a test person looks through a peephole with one eye at two white balls of equal size and distance in an otherwise dark black box ([Rock 1984](#) , 80f). In this experiment, all cues about the distance of the two balls coming from bottom-up information have been eliminated: there is neither (a) binocular depth information nor (b) monocular depth vision (e.g. by motion parallax via eye movement). If one tells the test person that one is a tennis ball and the other a ping pong ball, the tennis ball is perceived as being further away. This illustrates the utilization of acquired information when bottom-up information is missing. However, as soon as cues of sort (a) or (b) are admitted, for example, when the peephole is broadened, and the person is allowed to move her eye or to look with two eyes, the illusion disappears, and the person sees the two balls in equal size and distance.

Pro-5 : Scientific data (or observations i.w.s.) are theory-dependent. Scientists commonly use a wider concept of ‘observation’ or ‘datum’ than perception with the ‘naked eye.’ According to the scientific usage of this word, scientists observe galaxies, bacteria, or electrons with suitable *instruments*. A number of authors, including Kuhn in many passages, relate their thesis of theory-dependence to this observation concept i.w.s. [Giare \(1999 , 12f, 80\)](#) and [Kuipers \(2000 , sec. 13.3\)](#) argued that whenever a new measurement instrument is invented, a formerly theoretical concept is turned into an observational concept.

Contra-5 : Scientific data are theory-dependent interpretations of perceptions. The theories presupposed here are mainly theories about the way measurement instruments work, which are usually treated as an unproblematic part of the advanced knowledge of a scientist. This *pragmatically* extended observation concept i.w.s. is undoubtedly very important for the efficient progress of science, since one cannot always put everything simultaneously into question, and the investigation of scientific problems rests on large amounts of background knowledge that is highly corroborated and taken for granted. However, this fact does not imply what some philosophers of science wrongly infer from it, namely that there *are* no observations in the narrow sense (i.n.s.) in science, that is, direct *perceptions* (mainly visual perceptions). We claim theory-independence only for this narrow concept of observation.

Most of the time, scientists agree on their observations i.w.s., and there is no need for them to make their direct perceptions explicit. But when scientists disagree on what they have observed i.w.s., they will critically reflect their implicit theoretical assumptions that led them to infer different observations i.w.s. from their common perceptions. For example, a pulmonologist A may report to have seen pneumonia in the X-ray image of a patient’s chest. What he has seen directly is a certain blackening pattern that he interprets as pneumonia. If another pulmonologist B doubts this diagnosis, then what B doubts are not the visual abilities of A. Rather, A and B will discuss whether the blackening pattern which they *both* see is indicative of pneumonia or not. In conclusion, the epistemological significance of observations i.n.s. lies not in being permanently verbally reflected, but in serving as a possible level of theory-neutral agreement to which scientists can step back in the case of disagreement about observations i.w.s.

Pro-6 : There is a continuity between observational and theoretical concepts. Maxwell ([1962](#) , 7), Carnap ([1966](#) , 255f), and Hempel ([1966](#) , 78–84) pointed out that there is a continuous transition from observability with the naked eye to intermediate stages such as observability with glasses or a magnifying glass, until observability with an optical telescope or microscope, an electron microscope, or a radar screen. A number of authors inferred from this that any demarcation between perception and theory-dependent observation i.w.s. is arbitrary.

Contra-6 : That there exists a continuous transition in the extent to which concepts are theory-laden is true. But first , the fact that there exists a continuous transition between two features like for example black and white, does not by any means imply that there is no difference between black and white (cf. van Fraassen 1980, 214). So, the distinction between observation in the narrow and in the wide sense is significant even if it is continuous. *Second* , there are marked jumps in this transition: while observation i.n.s. does not require any learned background knowledge, the required background knowledge increases successively for the usage of the above-named instruments. As an illustration of this point, imagine what a native tribesman would do with these instruments, without having received instructions. Presumably, he would understand glasses, but

already optical microscopes and telescopes would hardly be understandable, and that's not to mention electronic microscopes or radar screens, as the tribesman has no clue about the background knowledge that is needed for an adequate handling of these instruments.

Pro-7 : Meaning holism makes observation concepts theory-dependent. Quine ([1951](#)) doubted the traditional analytic-synthetic distinction (see [sec. 3.11](#)). His alternative semantic view was a radical meaning holism, according to which the meaning of a concept is determined by all that we know about it. If this were true, then theory-neutral observation concepts could not exist, because in scientific theories, observation concepts get connected with theoretical concepts. According to meaning holism, this would imply that their meaning becomes loaded with theoretical statements. In this sense, Chalmers ([1986](#) , 36) argued that “water” is a theory-laden concept, because chemical theory connects it with the molecular structure H_2O .

Contra-7 : Radical meaning holism will be criticized in [sec. 3.11](#) as more or less *untenable* . Quine's inner conflict concerning observation concepts is revealed in (1992, 7f), where he attributes a “two-faced nature” to the sentence “this is water:” “holophrastically,” so Quine, the sentence is free of theory, but by its composite structure, the sentence contains theory. This sounds like an ad hoc maneuver to escape the untenable consequences of meaning holism. We argue that the meaning of observational concepts is fixed by processes of definition or ostension such as “this is water.” This meaning is *not* affected by theories, which connect observational with theoretical concepts. If this were not so, then one would have to say that a physicist and a layperson mean entirely different things when they point to a creek and say that water is bubbling there – but this seems too strange to be true.

Pro-8 : Observation concepts (as all other concepts) are language-dependent. The thesis of the language dependence of experience was made popular by the philosopher Wilhelm von Humboldt and by the anthropological and ethnological research of Sapir and Whorf (see [Kutschera 1976](#) , [sec. 4.1](#)). Many native tribes, for example, have different and less differentiated concepts of color than members of Western civilization ([Berlin/Kay 1999](#)). From this fact, cultural anthropologists inferred that these natives really see the colors in their environment differently from us. The other way around, Inuits (Eskimos) have reportedly many more concepts for sorts of snow than us (see [Whorf 1956](#) , 216; [Kutschera 1976](#) , 251f). Whorf derived from this his *linguistic relativity principle* , according to which we can only perceive what is describable, and in a sense also prescribed, by the concepts in our language.

Contra-8 : It must be admitted that different cultures, to a certain degree, develop different systems of concepts, in which they describe the perceptions of their different environments. But from this, it does not follow that their sense experience itself is language-dependent. We could only conclude this if members of these cultures were also not able to *learn* the observation concepts of other cultures by *ostensive training* , that is, by acts of pure ostension. This ostensive learning ability is present in all cases, however. The Zuni Indians can learn the concept of “orange,” for which they don't possess a word, through brief ostensive training – that is, through simple acts of pointing (“this is orange,” “this is not orange”). The Tiv and the Dani were able to learn many European color concepts in this way, for which there is no word in their mother tongue ([Garnham/Oakhill 1994](#) , 49–51; [Berlin/Kay 1999](#) , 24). Similarly, a European can learn the Inuit's words for different kinds of snow ostensively.

2.9.2 A Definition of Theory-Neutral Observation Concepts

In this section, we present our definition of theory-neutral observation concepts. We start from the insight that the defining criterion cannot come from logic, but must be an empirical-psychological one. Hence, the definition of “observation concept” will be relative to the sensorial capacities of humans. For example, the magnetic field is non-observable for humans, but observable for animals with magnetic sense organs (cf. [Carnap 1936](#) /1937, 454f; van Fraassen 1980, 56f).

Traditionally, an observational concept is characterized as a concept whose applicability to an observable object can be determined through mere perception (e.g. [Carnap 1936](#) /1937, 454f). To verify this characterization, we must rely on the *introspection* of individuals – but it is well known that many people indiscriminately conflate perception with theoretical interpretation in their supposed observations. For example, esotericists claim to observe “spiritual fields,” and 18th-century chemists reported observing

phlogiston. To avoid this problem, we base our definition on an experimentally testable criterion: the *ostensive learnability* of concepts. With this step, we do not only circumvent the problem of the unreliability of introspection, but also the influence of language. An *ostensive learning experiment* is based on nonverbal behavior and involves only three communicative acts that are culturally universal: acts of ostension (“this”), interrogative forms (“is this X?”), and words or gestures for “yes” and “no.”

It is essential for an *ostensive learning experiment* to work in a *nonverbal* way. Otherwise, the linguistic relativist could argue that the Zuni have learned to see the color “orange” only by learning the Western color language. Also, Quine says in several places that the characteristic property of observation sentences is their ostensive learnability, independently of acquired information (1960, 42; 1974, 37, 41f; 1992, 3f, 5f). But this thesis of Quine is in strong tension with Quine’s radical meaning holism (cf. pro-and-contra 8 of [sec. 2.9.1](#)). So, in spite of these similarities, Quine’s theory of observation sentences is essentially different from our account.

An ostensive learning experiment consists of a *training phase* and a *test phase*. In the training phase, the concept to be tested for observability is introduced to the test person (tp) with an *artificial word* “X” unknown to them, so that the tp does not associate *any* meaning with the artificial word. The concept referred to by “X” is illustrated to the tp with a set of positive and negative instances of the concept, and the information “this is an X” or “this is not an X” is given. The instances may be presented in the form of photos, movies, or real objects. They must be clearly visible by the tp; more precisely, the experiment takes place under ‘normal observation conditions’ (see below).

The ostensive learning process to be achieved by the tp consists in the abstraction of the meaning of “X,” that is, the property cluster corresponding to X, from the examples presented. Whether the tp has succeeded in this is investigated in the *test phase* which follows, in which a series of *new* positive and negative instances of the concept are presented, and the question “is this an X?” is asked. The tp’s success frequency in correct answers in the test phase indicates to which extent the tp has successfully abstracted from the training instances the perceptual property cluster that corresponds to the concept “X.”

If a concept can be learned from a small set of learning instances by (almost) all test persons, *independently* of their background culture and even if the tps possess no word for this concept in their spoken language(s), this is a more-or-less compelling reason to see this concept as a *theory-independent* observation concept. Without this assumption, the learning success would be a sheer miracle. If a Zuni Indian were unable to perceive the color feature “orange” because she has no concept of it, then she would not perceive the difference between the orange, red, and yellow pictures in the training phase, and it would be *inexplicable* how she could successfully pass the test phase. In conclusion, we propose the following definition of observation concept:

(Def. 2.9-1) A concept is a *theory-neutral observation concept* (or an observation concept i.n.s.) iff almost all humans can acquire this concept in an ostensive learning experiment, under normal observation conditions, independently of their background information, language, and culture.

Thereby, the *normal observation conditions* include (i) the normal physical conditions (sufficient light etc.), (ii) normal biological conditions (no biological deficiencies in the person’s sensory organs and nervous system), and (iii) normal psychological conditions (the person is in a wakeful and mentally normal state, and her yes/no utterances are truthful). All three conditions can be experimentally tested.

Our definition of observation concepts does not imply that all cultures have the same system of observation concepts (this is de facto false, as we know). What the definition implies is that different cultures can learn each other’s observational concepts in an ostensive manner, without being forced to adopt the language or worldview of the other culture. This consequence accords with the facts mentioned in [sec. 2.9.1](#) (contra-8), namely that native tribes can learn the Europeans’ color concepts, or that Europeans can learn the Inuit’s concepts for kinds of snow (etc.).

It would be interesting to perform ostensive learning experiments of the described sort (according to my knowledge, this has not been done so far). Generally speaking, the number of instances that are necessary to grasp the perceptual property cluster that corresponds to the concept “X” will not only depend on the question whether the property cluster is observable in principle or not, but also on the *complexity* of this property cluster. So, the result of an ostensive learning experiment is expected to be a *gradual learning curve*

. This turns the resulting notion of an observation concept into a *gradual* one: a concept may be more or less observational. Nevertheless, we expect that there will be clear statistical divide between the following types of concepts, as illustrated in [fig. 2.9-3](#) (cf. [Schurz 2015b](#)):

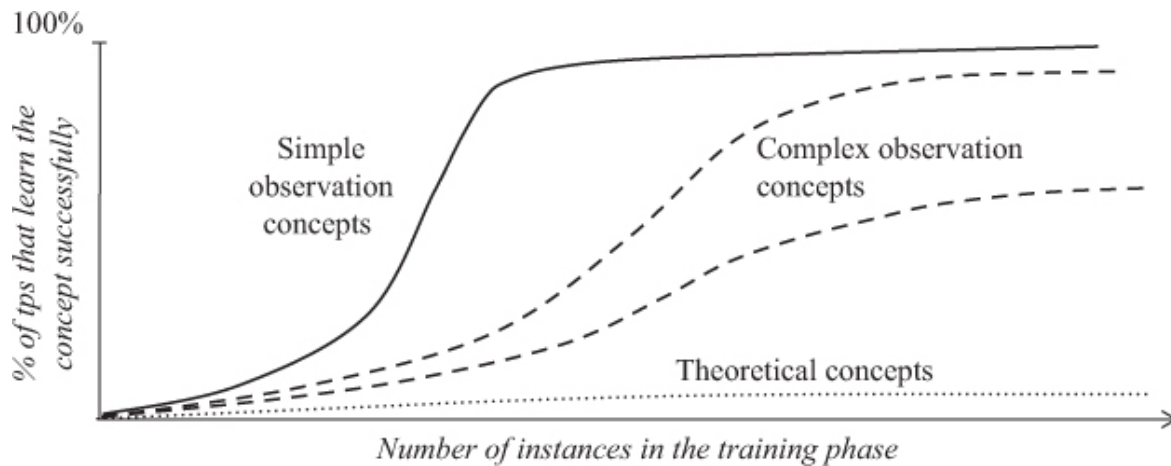


Figure 2.9-3 Expected learning curves in ostensive learning experiments. [↗](#)

1. For *simple* observation concepts (such as “red,” “square,” or “bird”), success frequencies will after a small number of instances climb above the significance threshold, toward values slightly smaller than 100% of tps.
2. For *complex* observation concepts (such as “either square and red, or circle and blue” or “hooded crow”), the success rate will climb slowly and will need many instances; often the success frequency will stagnate at lower levels, for example, at 50%.
3. For genuinely *theoretical* concepts (such as “atom,” “electric force,” or “oxidation”), the learning curve will not climb at all, even after many instances.

2.10 Value Neutrality in the Context of Contemporary Debates

2.10.1 The Role of Values in Science

As pointed out in [sec. 2.5.2](#), the value-neutrality requirement ([prop. 2.5-1](#)) possesses a critical-empiricist function. This stands in contrast to an influential division of epistemological paradigms according to their respective knowledge-guiding interests, which goes back to Habermas ([1966](#), 241ff) (cf., e.g., [Kriz et al. 1990](#), 151; [Rosenberg 2008](#), 127ff):

<i>Epistemological paradigm</i> :	<i>Knowledge interest:</i>
1 empirical-analytical	technical-instrumental controllability
2 hermeneutic	establishment of normative consensus
3 critical-dialectical paradigm	emancipation from power structures

Given our understanding of scientific value neutrality, these assignments are clearly incorrect. Science, being classified by Habermas as empirical-analytical, cannot be assigned to *any* specific interests that are external to science, but serves directly merely the science-internal epistemic goals, the search for content-rich truths. Scientific knowledge may be equally well used for technical purposes or for purposes of political reform. Adherence to value neutrality certainly does not mean that a scientist must refrain from practical involvement in society, or remain unaware of her social *responsibility* (cf. [Prim/Tilman 1979](#), 139f; [Koertge 2000](#), S48). All that the value-neutrality requirement implies is that the scientist *separates* her scientific knowledge from

fundamental value assumptions which she assumes in means-end inferences, and thereby *relativizes* her means-end inferences hypothetically to these value assumptions (recall the example in [sec. 2.5.2](#)).

The true difference between Habermas' three epistemological paradigms lies *not* in differing interests, but in the different treatment of the *value-neutrality question* (cf. [Schurz 1998a](#) , [2024b](#)). Representatives of the empirical-analytical program are (unfortunately) the only ones striving for a clear division between description and value judgments. In hermeneutic approaches, by contrast, descriptive statements and value judgments are often deliberately not distinguished from each other (cf. [Hilgendorf 2000](#) , 39ff; [Dupré 2007](#)), and in critical-dialectical approaches, certain emancipatory norms are presented in a quasi-scientific make-up.

Our analysis has straightforward consequences on some debates on the role of values in science ([Kincaid et al. 2007](#) ; [Carrier et al. 2008](#)). A significant number of contributors in this debate have criticized the ideal of value-free science. Their objections can be resolved by our analysis of means-end inference and the resulting requirement of separating empirically confirmed knowledge from value assumptions. Dupré ([2007](#) , 30–35), for example, argues that science needs to employ value-loaded terms in order to be practically relevant. Dupré is right insofar as the derived norms recommended by scientists employ value-loaded terms. But he seems to overlook that the scientist (qua scientist) recommends these derived norms only *hypothetically* and instrumentally, as means for certain fundamental norms that are given by society.

As emphasized in [prop. 2.5-1](#) , the requirement of value neutrality applies only to the scientific *context of justification* (CJ): it must be independent of fundamental, extra-scientific value assumptions. By contrast, the *context of origination* (CO) inevitably involves extra-scientific values that influence the selection of research questions. However, proponents of scientific value-involvement – such as Douglas ([2000](#)) – argue that even the CJ is influenced by extra-scientific values. In support of this position, the following argument is often made: First, the formulation of the research question in the CO is guided by non-epistemic interests. Second, in the CJ, the research question significantly shapes the choice of scientific method. This shows that non-epistemic interests can have an *indirect* influence on methodological choices, or more broadly, that choices within the CO can have an impact on the CJ.

The argument of indirect influence is certainly correct, but may create confusion. Once a research question is precisely specified, the choice of the optimal method for finding an insightful answer becomes a *purely epistemic* matter. As shown in Schurz ([2024b](#) , 164–167), this also applies to the determination of the “significance level” that is often-cited in this context. In Parker's ([2024](#)) words, the specification of the research question projects an initially non-epistemic interest into a purely epistemic research task. We illustrate this with an *example* : A research project on effective treatments for depression may be guided by non-epistemic interests of the pharmaceutical industry. As a result, the research question may be limited to pharmaceutical interventions: Which antidepressant drugs produce the best outcomes (operationalized in terms of efficiency, sustainability, etc.)? For a research question that has been specified in this precision, the choice of the optimal method (data collection, statistical procedures, etc.) no longer depends on non-epistemic interests. This holds true even if it turns out that, for certain types of depression, psychological therapies are more effective than antidepressant medication. In that case, a properly conducted study would simply yield a largely negative result – that is, antidepressants would be shown to have only limited effectiveness. ⁵

The example of pharmaceutical depression studies leads to another important insight. In our example, it becomes evident in the CJ that restricting the investigation of depression therapies to pharmacological treatments impedes research, because it neglects important causal factors that plausibly lie within the psychological domain. In such a case, the selection of research parameters initially made in the CO must remain open to later revision based on findings in the CJ, even if such revisions run counter to the extra-scientific goals of the research project. Otherwise, one can no longer speak of the CJ being independent from external value commitments.

As a final application of our concept of value neutrality, we turn to the so-called *new demarcation problem* ([Wilholt 2009](#) ; [Holman/Wilholt 2022](#)). This challenge arises for all opponents of value neutrality. The challenge consists in the fact that, on the one hand, advocates of value-laden science massively criticize the capitalist value bias in industry-sponsored studies, while at the same time maintaining that value influences on scientific knowledge are both inevitable and desirable. But this implies that they should regard not only their own (typically left-leaning) values as legitimate, but also those underlying industry-funded

capitalist research (cf. also [Henschen 2021](#) , 20). Most proponents of value-laden science seek to avoid this implication – but how? Some authors have proposed resolving the new demarcation problem by suggesting that scientists should rely on a specific set of preferred values – namely, those reflecting democratic majoritarian values forming the normative foundation of our society. ⁶

Le Bihan ([2024](#)) has raised a fundamental objection to the thesis of democratic value-dependence: modern democratic societies are polarized, with left- and right-leaning segments of the population holding opposing value preferences on many issues – from immigration to climate policy. In this context, the proposal to ground science in “democratic values” offers little guidance, as it leaves open on which side of the political spectrum these “democratic values” are supposed to lie. Moreover, the proposal carries a disturbing implication: in times of political upheaval – such as the current situation in the United States – major scientific findings and textbooks would have to be rewritten, since they would depend on democratic majority values that have now fundamentally shifted. Fortunately, the prevailing conception of science in modern democratic societies has largely freed itself from such political entanglements – and it should remain that way.

From a value-neutral perspective, the question of the “right values” is not a fundamental problem, even in times of value polarization. This is because the goals to which scientists relate their recommendations for means are determined by society. If society is in a state of *value polarization* , then value-neutral policy advice requires conducting hypothetical cost-benefit analyses for *multiple, diverging* value preferences that reflect the political positions of relevant segments of the population (cf. [Carrier 2022](#)). This was also a key element in the learning process reflected in the recommendations of the German Leopoldina commissions during the COVID-19 pandemic (see [Wiesing et al. 2021](#) ; [Schurz 2024b](#) , 158f).

2.10.2 The Argument of Inductive Risk

A much-debated objection against scientific value neutrality is the argument from inductive risk (AIR), originally proposed by Rudner ([1953](#)) and further developed by Douglas ([2000](#)). The AIR assumes that scientific hypotheses are typically supported by empirical evidence only with a certain degree of probability. When scientists accept a hypothesis H as true, they thereby incur a certain “inductive risk” – the risk of error. How small must H’s risk of error be in order for it to be reasonable to accept H as true – that is, to rely on H and be willing to act on the basis of H? The *risk* associated with accepting a hypothesis H depends on two factors: (i) the *error probability* of H (i.e. the probability of $\neg H$ given the total evidence) and (ii) the practical *cost* incurred by accepting H when H is false. While (i) is an epistemic, (ii) is a non-epistemic question, depending on extra-scientific values. H’s total risk of error is given as the product of (i) and (ii). It follows that for hypotheses with practical relevance, their acceptability as a basis for practical action depends inevitably on extra-scientific values.

Quite a few authors see the AIR as convincing evidence for the influence of non-epistemic values on the CJ ([Holman/Wilholt 2022](#) , [Brown 2024](#)). In reality, however, the AIR is weaker than it sounds – a “storm in a teacup.” The AIR can be defused by the demand of *uncertainty transparency* in science, that is, by explicitly stating the error probabilities of scientific claims. This demand not only serves the goal of value neutrality but also helps prevent problematic consequences of expertocracy. We illustrate this with the example of the prediction of *earthquake safety* . Let H be the hypothesis that (in a given region over the next few days) no earthquake of magnitude greater than 6.5 on the Richter scale will occur. Even if the error probability of this hypothesis is only 5%, many affected people would be poorly advised to act on it – at least those whose lives would be seriously endangered in case of error, because their houses are poorly built and at risk of collapse at this level of seismic activity. It is therefore undoubtedly correct that the question of whether to accept non-H rather than H – and thus whether or not to evacuate – depends on non-epistemic cost-benefit considerations (for detailed calculation s. [Schurz 2024b](#) , 160–167 and [sec. 5.9.5](#)).

But does this really mean that scientists should make such value judgments on behalf of the affected people, as proponents of the AIR would have us believe? Certainly not. Instead, the error risks involved should be explicitly stated and explained. The L’Aquila earthquake illustrated this in a tragic way: In October 2012, six earthquake experts and a government official in Italy were convicted in the first instance and sentenced to several years in prison because they had failed to predict an earthquake that occurred on April 6, 2009, in L’Aquila, in which more than 300 people died – instead, they had given the all-clear ([Cocco et al.](#)

[2015](#)). The seismologists defended themselves against the verdict by pointing out that they had not issued a categorical all-clear, but rather a more cautious probabilistic judgment, namely that the observed slight increase in seismic activity was *not significant* , with an estimated error probability of 5%. As a result, the verdict was overturned in 2014; only the government official who had communicated the “all-clear” message to the media was ultimately convicted.

The AIR is based on the mistaken belief that it is the task of science to provide the public with categorical knowledge claims. Quite the contrary: objectively value-neutral science requires that for practical uncertain knowledge claims, the estimated error probabilities should always be reported – an insight first emphasized by Jeffrey ([1956](#) , 237). Referring to Schurz ([2013](#)), both Betz ([2017](#)) and Henschen ([2021](#)) have likewise proposed to solve the problem of inductive risk by making error probabilities explicit. Douglas and like-minded authors attempt to refute this position by arguing that people expect statements from the “authority of science” with clear-cut practical implications – such as giving an “all-clear” ([Douglas 2000](#) , 563; [2009](#) , 135; [Wilholt 2009](#) , 94). But scientists must not give in to such false expectations; rather, it is part of their responsibility to enlighten the public about the limits of the knowable and the risk of error.

However, the critics are right to point out that a statistical significance statement, hardly understandable for laypeople, is not sufficient for such public education. We suggest that error probabilities should be made comprehensible to non-experts by means of *hypothetical cost-benefit considerations* . In the earthquake example, the 5% error probability could be explained as follows: *If* the negative consequences of an earthquake occurring while one remains in their home are estimated to be more than 20 times greater than the costs of a temporary evacuation, *then* it would be advisable to evacuate.

2.10.3 Meta-ethical Justification of Value Neutrality and the Difference between Experience and Valuation

The value-neutrality requirement is based on the thesis that fundamental norm and value sentences are not capable of being justified by empirical science. In these subsections, I discuss this thesis in the context of *meta-ethics* , which is that part of ethics that investigates the meaning and justifiability of ethical statements. First and foremost, value neutrality requires the separation of descriptive statements from value and normative judgments. In formal scientific languages, this separation is already ensured by the differing forms of expressions. However, in everyday language, a problem arises: many terms in our everyday vocabulary carry *both* factual and evaluative content, often closely intertwined, making conceptual separation difficult. Some critics of value neutrality, such as Putnam ([2002](#) , ch. 2) and Dupré ([2007](#)), have claimed that in the case of so-called thick concepts such as “cruelty” or “rape,” such separation is impossible. Yet, as Hare ([1981](#) , 74f) has already argued, this separation is more a *matter of will* than of ability, and with the help of suitable linguistic tools, it is always possible.

The standard meta-ethical justification for the value-neutrality requirement is the *fact – value* or *is – ought* dichotomy , which goes back to David Hume and asserts that value or normative statements cannot be logically or analytically derived from purely descriptive statements ([Hume 1739](#) /1740, 177f). Attempts to prove this prima facie highly plausible thesis by means of modern logic were confronted with unexpected problems, such as the paradox of Prior ([1960](#)) (cf. [Stegmüller 1979a](#) , 183–186). As a result, the dichotomy thesis was called into question by a number of philosophers ([MacIntyre 1981](#) ; [Putnam 2002](#)) and could only be placed on a solid foundation through more recent logical investigations. ⁷ Schurz ([1997](#)) proves that it is only possible to derive nontrivial ethical statements from descriptive premises if one already presupposes so-called *is-ought bridge principles* in one’s ethical theory, such as the utilitarian principle “Whatever is desired by the majority of humans, is ethically good, or ought to be realized.” Is-ought bridge principles are not analytically valid, but are synthetic assumptions belonging to one’s accepted ethical system and are typically controversial. Therefore, norms or values cannot be justified by descriptive (non-ethical) premises without already assuming hidden ethical principles.

There exists a variety of different meta-ethical positions (e.g. utilitarianism and a priorism), but none of them stands in a substantial conflict with Hume’s Is-ought thesis. ⁸ The only exception is the position of so-called *value empiricism* (cf. [Czaniera 2001](#) , ch.1–2). This position assumes a special kind of “value experience,” which, although distinct from biological sensory experience, is analogous to it in that it relates to individual experiences and can serve as a basis for testing in the value sciences. This position has been

advocated by, among others, Firth ([1952](#)), Weingartner ([1978](#) , §7.15 ff), and more recently by Anderson ([2004](#)). In the following, I will use two arguments to show that such “value experiences” behave quite differently from genuine sensory experiences: the normal conditions argument and the independence argument.

A The *normal conditions argument* says that intersubjectively deviating perceptual judgments can almost always be traced back to *defects* in the normal conditions of the observation. But this does almost never apply to intersubjectively deviating value judgments. Let us consider, for example, the presumably true observation sentence: “There is a fire there.” The thesis of the *intersubjectivity* of such an observation sentence implies that any or almost any person would agree with this sentence, provided the *normal conditions* of observation explained in [sec. 2.9.2](#) are fulfilled. If a person within visual range of the fire would really repeatedly and persistently claim “but I see no fire,” then we would have every reason to assume that there is a defect in the person’s normal conditions, and we would be able to find out this defect by independent tests. Such a person would most probably be blind or brain-damaged, or could be in a state of extreme hallucination. For this reason, we would also be entitled to try to *cure* such a person of their observational defect.

But now let us assume that an adult person, under the same normal conditions of observation, systematically *disagrees* with a basic value sentence such as “It is good that *this* person is being kept from dying by being fed,” although we assume that 99% of all persons do agree with it. Then we would *not be entitled* to assume that there is a defect in the ‘normal moral condition’ of this person, that is, that this person is morally inferior and should be cured, just like the blind should be cured of their blindness. Rather, we must *respect* this person’s position *prima facie* as a deviating value position and also – as far as it does not infringe the rights of others – accept it. It is normal in Inuit society, for example, not to give food to a person who is so weak with age that they cannot feed themselves, but to let them die instead, because this is seen as a natural sign that their spirit wants to leave this world. Should we therefore declare the morals of the Inuit to be ‘morally defective,’ and try to ‘cure’ them of their defects? If we treated basic value sentences exactly like observation sentences, we could find ourselves moving dangerously close to medieval religious moral practices.

I do not want to deny with this thesis that there are indeed cases of *psychopathological* blindness of conscience, which are in need of a psychological cure – but these cases are not the *rule*. An extreme claim such as “I love killing other people” would rightly cause anybody to react by thinking “this person has a defect” – but this situation applies only to *very few* deviating basic value sentences, whereby the analogous situation in the case of deviating judgments of perception is the *rule*. Quite in this sense, Prim and Tilman ([1979](#) , 93f) report an experiment on the objectivity of the observation sentence: “there is a piece of chalk in front of me,” which they conducted on their students: the students reacted very indignantly, and declared that dissent in such questions was a problem for an *optician*, and not one of philosophy.

B The *independence argument* goes back to Toulmin ([1950](#) , 127) and runs as follows: Our ‘value experiences’ are *dependent* on our *moral background assumptions*. Our perceptual impressions, on the other hand, remain the same, even when our descriptive background assumptions change radically. For example, we see the Sun make its way across the sky day in, day out. Let us assume that up to now we have held the geocentric view (the Sun revolves around the stationary Earth); but now, after detailed study of Kepler and Newton, we have changed to the heliocentric view (the Earth revolves around the stationary Sun). There has now been a radical change in our background theory – but as already explained in [sec. 2.9.1](#) (contra-3), our visual experience is changed *in no way* by this fact: just as before, we see the Sun drawing an arch across the sky every day. Our visual experience is independent of acquired background assumptions.

This is quite different in the case of value-based moral sentiments. If I see somebody donating a sum of money to the poor, I may spontaneously have the moral sentiment: *that was a good deed*. If I find out, however, that the donor is a politician about to stand for election, who has invested this donation as part of a calculated plan to win votes, but is otherwise more concerned about the well-being of the rich than of the poor, then a quite different moral sentiment will arise within me: I now see the politician as a *hypocrite*.

The normal conditions and the independence arguments strengthen our thesis that an empirical justification of value or norm sentences in analogy to empirical science is not possible. Value and norm-

based judgments are not based on sensory experiences, but on subjective interpretations of experiences that depend on cultural background values and emotional impressions arising in the context of the perceived situation.

2.11 The Justification of Induction: An Unsolvable Problem?

2.11.1 Is Induction Rationally Unjustifiable?

As mentioned in [sec. 2.7.1](#), Hume's problem of induction consists in the question of how inductive reasoning can be rationally justified. David Hume made it clear to the philosophers of his time that, *first*, inductive inferences cannot be justified by means of *logic*, because it is logically possible that the future (or the unobserved) is radically different from the past (or from the observed). However, Hume went much further: he cast doubt on the rational justifiability of induction as a whole (1748, sec. 4). Induction, *second*, also cannot be justified by observation, because in inductive inferences one infers that which has *not yet* been observed. And *third*, inductive inference also cannot be justified by induction, as this would lead us straight into a *circular argument*: if we think inductive inference is reasonable because it has been corroborated so far, then we are assuming that what has been corroborated so far will – by inductive inference – also be corroborated in the future. In such a circular argument, we assume exactly the thing we want to show, and so we show nothing at all. All three traditional methods of scientific justification – logic, experience, and induction – thus appear to fail when it comes to justifying induction itself. Hume was also the first to show that, *fourth*, a reformulation of the induction problem using probability theory does not help (1748, sect. 6). If we argue that inductive inferences do not always, but at least *with great probability* lead us from true premises to true conclusions, then we are assuming a probabilistic induction principle, according to which the *event frequencies* observed so far may be inductively transferred to the future or to non-observed cases.

These were the reasons which led Hume to his deeply *skeptical* thesis, that inductive inference is not capable of being rationally justified, but is based merely on psychological *habit* (1748, sect. 5). Finding a satisfactory solution to Hume's induction problem would be of the greatest importance. And yet, to this day, there is no generally accepted solution to the problem. ⁹We briefly discuss five solution attempts that have been suggested in the last few decades but could not find general acceptance:

1 *Induction is not needed in science?* Popper's attempt to circumvent the induction program was to argue that scientific methodology doesn't need induction. In [sec. 2.7.2](#), we have seen, however, that Popper's argument fails, because Popper's own methodology of "corroboration" contains in its heart the principle of epistemic induction.

2 *Induction as rational by definition?* Proponents of the analytic account (e.g. [Strawson 1952](#), 257; [Ayer 1956](#), 74f, [Pollock 1974](#), 204) have argued that induction is rational by the very definition or meaning of rationality. Recall, however, that by a rational justification we don't mean a mere 'intuition of rationality' (which is then declared as being 'true by definition'), but a demonstration that inductive inferences are truth-conducive, in the sense of optimizing our success in delivering true predictions. This cannot be shown by a mere 'definition of rationality,' no more than diseases can be cured by a 'definition of health.' Nor is Strawson's claim (ibid., 249) correct that we cannot imagine a situation in which inductive inferences fail (cf. [Salmon 1974](#)). We can not only conceive of such situations in thought (consider *Alice in Wonderland*); in fact, millions of people *do* believe in the superiority and even in the 'higher rationality' of non-inductive 'methods,' be it clairvoyance, fortune-telling, or God-guided inner intuition. So, a satisfactory justification of induction would not only be of high philosophical, but also of high *cultural* importance, as part of the enterprise of explaining scientific rationality.

3 *Rule-circular 'justification' of induction by induction?* The circularity involved in the 'inductive justification of induction' has been called *rule-circularity* ([Ladyman/Ross 2007](#), 75), because here, the conclusion (the reliability of induction) is not an element of the premise set, but is presupposed by the underlying inference rule (namely induction). Several philosophers expressed the hope that rule-circular arguments have epistemic value and might offer a solution to the problem of induction. ¹⁰The following counterexample by Salmon ([1957](#), 46) destroys this hope. It shows that the same type of rule-circular

argument that ‘justifies’ the reliability of induction can be used to ‘justify’ the rule of counterinduction, which says, roughly speaking, that one should predict the opposite of what one has observed:

(2.11-1)

Rule-circular justification of induction :

Premise : Past inductions have been successful.

Therefore, by rule of induction:

Inductions will be successful in the future

Rule-circular justification of counterinduction :

Premise : Past counterinductions have not been successful.

Therefore, by rule of counterinduction:

Counterinductions will be successful the future.

The two circular arguments share exactly the same structure, the premises of both arguments are true, and yet they lead to opposite conclusions. This proves that such rule-circular arguments are epistemically worthless pseudo-justifications. A generalization of this refutation for inferences to the best explanation is found in Douven ([2011](#)), and for arbitrary rule-circular arguments in Schurz ([2024a](#) , [sec. 4.3](#)). Rule-circular justifications can also be given for dangerous fundamentalistic rules, such as the “rule of blind trust in my leader,” abbreviated (BTL), whose rule-circular justification goes as follows: “I believe in the reliability of the rule BTL because this was assured to me by my leader.”

4 Induction justified by uniformity of nature? John Stuart Mill ([1865](#) , III .3.1) and Bertrand Russell ([1912](#)) held the view that the reliability of induction could be justified by assuming the uniformity of nature (i.e. nature’s future resembles its past). The main problem with this approach is that it also leads to a *justificatory circle* , since the assumption of uniformity can itself only be justified by an inductive inference. Note that as in ([2.11-1](#)) above one may give a corresponding circular justification of counterinduction by counteruniformity: counterinduction is justified by counteruniformity which is in turn justified by counterinduction from past observations. A “localized” version of the uniformity-account has been proposed by Norton ([2003](#)). For Norton, inductive reasoning is not governed by formal but by domain-specific or “local” inference rules such as “if some samples of Bismuth melt at 271 ° C, then all samples of Bismuth will do so.” Local inductions are in turn justified by local uniformity assumptions, for example, “samples of the same element agree in their physical properties.” However, Norton’s proposal is subject to exactly the same charge of circularity as the Mill-Russell approach, since local uniformity assumptions must be justified by inductive inferences (cf. [Kelly 2010](#) ; [Schurz/Thorn 2020](#)).

5 Induction justified by inference to the best explanation? Several authors ([Harman 1965](#) ; [Armstrong 1983](#) , sec. 6; [Lipton 1991](#) , 69) have argued that inductive inferences are justified because they are instances of inferences to the best explanation (IBE). Recall from [sec. 2.7.3](#) that an IBE infers from an observed phenomenon that the premises of its best (available) explanation are probably true. An inductive generalization – so the argument goes – is an instance of an IBE , because the best explanation of a regularity observed so far (e.g. “so far all observed ravens were black”) is the corresponding lawlike generalization (“It is a law that all ravens are black”).

The major problem of the IBE-account is, however, that the justification of IBEs is even more difficult than the justification of induction. This is so because IBEs involve even more uncertainties than inductive inferences (recall [sec. 2.7.3](#)). Analogous justification strategies to those discussed above for induction have been proposed for IBEs (cf. [Bird 1998](#) , 171). Armstrong ([1983](#) , 53), for example, has argued that IBE is rational ‘by definition.’ Lipton ([1991](#) , 167ff) and Psillos ([1999](#) , 82) suggested to justify IBE in a rule-circular way. But we have seen that both strategies are bound to failure. Moreover, it is unclear why one should regard the explanation of “All so-far observed Fs have been Gs” by “Because it’s a law that all Fs are Gs” as a *best* explanation, if one does not accept the inductive premise that the world is uniform. If one assumes instead that from time to time the laws of nature undergo radical changes, the explanation “Because it was *so-far* a law that Fs are Gs” seems to be the better explanation.

2.11.2 The Optimality of Meta-induction

Due to the many failed attempts to solve the problem of induction, quite a few contemporary philosophers consider it unsolvable. ¹¹ We do not share this view. We agree with Hume's skeptical diagnosis that no noncircular argument can establish that induction is *reliable*, that is, will be predictively highly successful, whatever the future looks like. However, there is an alternative justification strategy that is not blocked by Hume's argument: namely, the attempt to show that induction is an epistemically *optimal* prediction method, in the sense of being most successful among all prediction methods that are *accessible* to a given agent. Even in radically skeptical scenarios in which all methods of prediction are unreliable, induction can still be optimal in the sense of 'being the best of a bad lot.'

A noncircular optimality justification of induction was developed by Schurz (see, e.g., [Schurz 2008b](#), [2019a](#)). A precursor to this approach is Hans Reichenbach's justification of induction as the "best alternative" ([Reichenbach 1949](#), §91). However, Reichenbach failed to develop a robust optimality argument in favor of induction, as he based his argument on *object-induction*, that is, induction aimed at predicting events. Yet, nothing can rule out the logical possibility of a clairvoyant who predicts a random sequence with a significantly higher success rate than the scientific inductivist (see [Skyrms 1999](#), sec. III.4). By contrast, the approach developed by Schurz is based on the idea of *meta-induction*.

The meta-inductivist applies the inductive method to all accessible prediction methods and attempts to derive an optimal prediction strategy based on their observed track records. The simplest meta-inductive method is *Imitate-the-best*, abbreviated as ITB. It predicts what the actually best non-MI method predicts (resolving ties according to a given ordering of accessible methods). ITB's success will be optimal in the long run if, after some finite time, there is a *unique* best method that ITB imitates forever (cf. [Schurz 2019a](#), [sec. 6.1](#)). Interestingly, ITB is *not* universally optimal. ITB's optimality breaks down when it plays against 'deceiving' methods that lower their success rate as soon as their predictions are imitated ([Schurz, 2008b](#), sec. 4). A realistic example is predictions of stock values in a 'bubble economy:' here, the prediction that a given stock will yield a high return leads many investors to buy that stock, causing it to crash by doing so (since the stock lacks sufficient economic support).

A more refined meta-inductive strategy is attractivity-weighted meta-induction. This meta-inductive method, abbreviated as AW, predicts a weighted average of the predictions of all methods accessible to it, with weights determined by the methods' past success rates (overview in [Schurz 2019a](#), sec. 6–7). Building on results from mathematical learning theory ([Cesa-Bianchi and Lugosi 2006](#)), Schurz ([2008b](#)) could show that AW is indeed universally optimal: in all prediction games with arbitrary sequences of events and a finite number of available methods, AW's predictive long-run success is greater than or equal to the maximal long-run success of all accessible methods ([Schurz 2008b](#), theorem 4; [2019a](#), theorems 6.8–6.11). Even in 'radically nonuniform' scenarios in which the success rates of the leading players permanently oscillate around each other, AW's success rate converges to the maximal success rate (or to a greater value), with a small short-run loss that converges to zero with an increasing number of predictions.

The precise formulation of the meta-inductive optimality theorem runs as follows. Let $(e) =_{\text{def}} (e_1, e_2, \dots)$ be a sequence of events e_n ordered along a discrete time $n = 1, 2, \dots$, whose possible values are coded by real numbers in the interval $[0,1]$. Let $M = \{M_1, \dots, M_m, AW\}$ be a finite set of (m) simultaneously accessible prediction methods together with AW. The methods $M_1, \dots, M_m$ may be given as computational algorithms, real-life experts or even (purported or real) clairvoyants; their "simultaneous accessibility" means that the epistemic agent can track and compare their success records, either by observation or by simulation. The pair $((e), M)$ is called a *prediction game*. In each round n of the game, it is the task of each method to deliver a prediction $\text{pred}_{n+1} \in [0,1]$ of the next event e_{n+1} of the sequence. The deviation of the prediction pred_n from the event e_n is measured by a normalized loss function, $\text{loss}(\text{pred}_n, e_n) \in [0,1]$. The optimality theorem below holds for all *convex* loss functions (a generalization is possible that even holds for all loss functions). The predictive success of a method M at a given time point n is evaluated by its *success rate* $\text{suc}_n(M)$, defined as the sum of M 's scores achieved until time n divided by n (where M 's score achieved at time n is $1 - \text{loss}(\text{pred}_n, e_n)$). The attractivity of a method M for AW at time n , $\text{at}_n(M)$, is defined as $\text{suc}_n(M) - \text{suc}_n(AW)$. The prediction of the meta-inductive method AW for the next time $n+1$, $\text{pred}_{n+1}(AW)$, is the normalized weighted average of the predictions of the methods $M_1, \dots, M_m$ for this time with weights

defined as $w_n(M_i) =_{\text{def}} e^{\eta \cdot n \cdot \text{at}_n(M_i)}$, where $\eta = \sqrt{8 \cdot \ln(m)/(n+1)}$. “ maxsuc_n ” denotes the maximal success rate of the accessible methods at time n . With these definitions, the optimality theorem runs as follows:

(Prop. 2.11-1) *Universal optimality of AW*: ¹²

For every prediction game $((e), \{M_1, \dots, M_m, AW\})$:

- i. (Short run) $(\forall n \geq 1:) \text{suc}_n(AW) \geq \text{maxsuc}_n - 1.77 \cdot \sqrt{\ln(m)/n}$.
- ii. (Long run) $\lim_{n \rightarrow \infty} (\text{suc}_n(AW) - \text{maxsuc}_n) \geq 0$.

In the short run, weighted meta-induction may suffer from a possible loss (or “regret”) compared to the leading method, caused by the fact that AW must base her prediction for the next time on the *past* success rates of the accessible methods. However, proposition 2.11-1(i) states a worst-case upper bound for this loss, which is small if the number of rounds n is not smaller than the number of candidate methods m , and which quickly converges to zero when n grows large.

AW cannot get deceived by a method because it never favors only one method but a weighted average of all attractive methods. A second important feature of AW is that it tends to *ignore* methods that are constantly worse than itself, since exponential weights with negative attractivities quickly converge to zero. This ignorance-tendency is necessary for AW’s long-run optimality and a crucial distinction of AW compared to success-weighted meta-induction that uses simple success rates as weights and is provably nonoptimal (Schurz 2019a, sec. 6.8.1). It follows that in games in which one method dominates its competitors, the weight of the best method will quickly converge to 1, and AW’s behavior will converge to that of ITB.

The optimality result holds independently of any assumptions about the possible worlds or their prior probabilities. This means that meta-induction possesses a mathematical a priori justification. But what does this result imply for the justification of object-induction? In possible worlds that contain genuine clairvoyants, object-induction cannot be optimal; in such worlds, meta-induction does not favor the object-inductivists, but rather the clairvoyants. However, the a priori justification of meta-induction provides us with the following noncircular *a posteriori justification of object-induction*: we know from experience that object-inductive prediction methods have so far outperformed all other available prediction methods. Therefore, we are justified – by meta-induction – in continuing to use object-induction in the future. This justification is, unlike the “inductive justification of induction,” no longer circular, because we have an independent a priori justification of meta-induction.

In the meantime, a lively debate has emerged regarding the prospects of the meta-inductive optimality justification, in which further challenges have been raised and solutions proposed. ¹³In Schurz (2024a), optimality justifications are extended to the entire field of epistemology.

2.12 Introductory Literature and Exercises

2.12.1 Introductory Literature

On the method of rational reconstruction, see Stegmüller (1969 ff, vol. IV.1). A classic on the German value judgment debate is Albert and Topitsch (1971, ed.), for more recent debates, see Kincaid et al. (2007) and Carrier et al. (2008); a recent introduction is Elliott (2022). A collection on the question of the unity of sciences is Rahman et al. (eds., 2004). Psillos (1999) offers an excellent introduction to the empiricism versus realism debate on philosophy of science. Recent overviews on the metaphysics of science are Schrenk (2017) and Hüttemann (2021). Rock (1984) introduces into the cognitive psychology of perception. Swinburne (1974) and Henderson (2018a) offer introductions to the problems of induction. Lipton (1991) is a classic on inference to the best explanation and Niiniluoto (2018) is a good introduction to abduction in science.

2.12.2 Questions and Exercises

Sec. 2.1

1. What is meant with “context of discovery” versus “context of justification?”
2. Explain the normativist argumentation and its criticism.
3. What is the supreme epistemic goal of science? Does it presuppose metaphysical realism?
4. Explain the method of rational reconstruction.

Sec. 2.2–2.4

1. Explain the five common epistemological assumptions of the sciences.
2. In which sense are minimal realism and minimal empiricism “minimal?”
3. What is the difference between falsifiability and fallibility?
4. Explain the common methodological features of sciences.
5. To what extent are the concepts of “natural sciences,” “human and social sciences,” and “humanities” ambiguous?
6. How do empirical, experimental, and dissecting sciences differ from one another in regard to the degree of their empirical invasiveness? Classify educational science, human medicine, and cultural studies within this classification.

Sec. 2.5–2.6

1. Explain the means-end inference using examples from your area of study.
2. Explain the suggested formulation of the value-neutrality thesis and all concepts contained in it.
3. Discuss: How is the requirement of “value neutrality” related to that of “political correctness?” Are the two in conflict?
4. Discuss physical reductionism and its criticism. What is “supervenience?”

Sec. 2.7

1. Which kind of inference (deductive, inductive, or abductive) is being used in the following examples?
(i) Most light switches are working. I am pressing the switch. Therefore, the light will go on. (ii) All humans are mortal. Aristotle is human. Therefore, Aristotle too is mortal. (iii) My refrigerator has always worked well up to now. Therefore, I will be able to rely on it tomorrow. (iv) Every time Gregor talks about his brother, his facial expression becomes strained. Therefore, he has a problem with his brother.
2. Explain the principles of methodological, logical, and epistemic induction. What are the Popperian arguments against the first two kinds of induction?
3. Give an example of an inference to the best explanation.
4. Which inferences in 1 are monotonic and which are non-monotonic?

Sec. 2.8–2.9

1. What is the difference between metaphysical and hypothetical-constructive realism?
2. Explain the argument of radical constructivism against realism and its criticism.
3. One person could fail to notice another because (a) he is deliberately ignoring her, (b) he is blind, (c) he is asleep, or (d) it is pitch dark. Which normal condition of observation is infringed in each case?

Sec. 2.10–2.11

1. Explain Habermas’ threefold division of scientific disciplines according to their respective interests. How is this view criticized?
2. Discuss the normal conditions and the independence argument.
3. Recapitulate David Hume’s four arguments against the possibility of justifying induction.

4. Explain the structural equality of the inductive justification of induction and the counterinductive justification of counterinduction.
5. What is a meta-inductive prediction method? In which sense is attractivity-weighted meta-induction optimal?

Notes

1. Cf. Oppenheim/Putnam ([1958](#)), Causey ([1977](#)), Hofer/Smeenk ([2016](#)); Hüttemann ([2021](#) , ch. 5–6). [↗](#)
2. Cf. Humphreys ([1997](#) , 10), Ladyman/Ross ([2007](#) , sec. 1.6), and Hüttemann ([2021](#) , 151–155). [↗](#)
3. For arguments in this direction cf. Chalmers ([1986](#) , [sec. 3.4](#)), van Fraassen ([1980](#) , 56ff), Albert ([1985](#) , sec. II.8), Brewer/Lambert ([2001](#) , S182f). [↗](#)
4. For example, Putnam ([1962](#)), Shapere ([1982](#)), Kuipers ([2000](#)), Giere ([1999](#)), Adam ([2004](#)). [↗](#)
5. It is well known that the so-called *positivity bias* in scientific research often leads to the non-publication of negative results, which causes significant distortions in the research record ([Ioannidis 2005](#) ; [Schooler 2011](#)). One example is evidence of bias in pharmaceutical studies in favor of their industrial sponsors ([Brown 2008](#)). However, such distortions are not an inherent feature of science but rather a *defect* of scientific practice under external pressures. In response, a number of scientific journals have introduced measures to ensure the publication of negative research findings. [↗](#)
6. For example, Elliott ([2017](#)), Schroeder ([2021](#)), Alexandrova/Fabian ([2022](#)). [↗](#)
7. S. Pigden ([1989](#)), Schurz ([1997a](#)), Russell ([2022](#)); overviews in Hudson ([1969](#)) and Pigden ([2010](#)). [↗](#)
8. Overview of meta-ethical positions is, for example, found in Frankena ([1963](#)), Kutschera ([1982](#) , sec. 2), Schurz ([1997a](#) , sec. 11.3), and Sayre-McCord ([2023](#)). [↗](#)
9. Cf. Stegmüller ([1977](#) , vol.II, ch. 4), Swinburne ([1974](#)), Rescher ([1980](#)), Howson ([2000](#)), Ladyman ([2002](#) , ch. 1), Schurz ([2019a](#) , ch. 2–4). [↗](#)
10. For example, Braithwaite ([1974](#)), Black ([1974](#)), van Cleve ([1984](#)), Papineau ([1993](#) , sec. 5), Goldman ([1999](#) , 85), and Psillos ([1999](#) , 82). [↗](#)
11. See, e.g. Stegmüller ([1977](#) , vol. II, ch. 4), Greco ([2005](#) , 265); Williamson ([2023](#) , 451). [↗](#)
12. The loss term in (i) simplifies the loss term in th. 6.9(ii) of Schurz ([2019a](#)), since $\sqrt{2} + 1/\sqrt{8} < 1.77$. [↗](#)
13. See, for example, Arnold ([2010](#)), Henderson ([2018a](#) , sec. 7.3), Sterkenburg ([2019](#)), Henderson ([2023](#)), Douven ([2023](#)), Briesen ([2026](#)), Shogenji ([2023](#)), Wolpert ([2023](#)), Pitts ([2023](#)), Kimling ([2023](#)); Replies in Schurz ([2021b](#) , [2023a](#) , 2023b, 2024c, 2025). [↗](#)

3 The Conceptual Toolkit

Language, Logic, and Probability

DOI: [10.4324/9781003675075-3](https://doi.org/10.4324/9781003675075-3)

This chapter has been made available under a CC BY-SA license.

3.1 Kinds and Classification of Concepts

The basic logical building blocks of any language are its concepts. A concept is understood here as a linguistic expression *together* with its meaning. Grammatically, concepts are *beneath* the level of the sentence – they are *subsential*. “Grass” and “green” are concepts; “grass is green” is a sentence. Though concepts signify something, they cannot in themselves be true or false; only sentences composed of concepts can. Concepts may be *atomic* or *complex*; for example, “grass” is an atomic, while “green grass” is complex. Sentences are composed of concepts, and the nature of a sentence depends on the nature of its concepts. So, we will *first* deal with the various different kinds of scientific concepts, before dealing with the various different kinds of scientific sentences. Concepts may be classified in three respects: first according to their logical type, second according to their content type, and third according to their gradation (or scale) type.

3.1.1 The Logical Type of Concepts

Logical grammar classifies concepts according to their logical-semantic function. [Fig. 3.1-1](#) presents an overview.

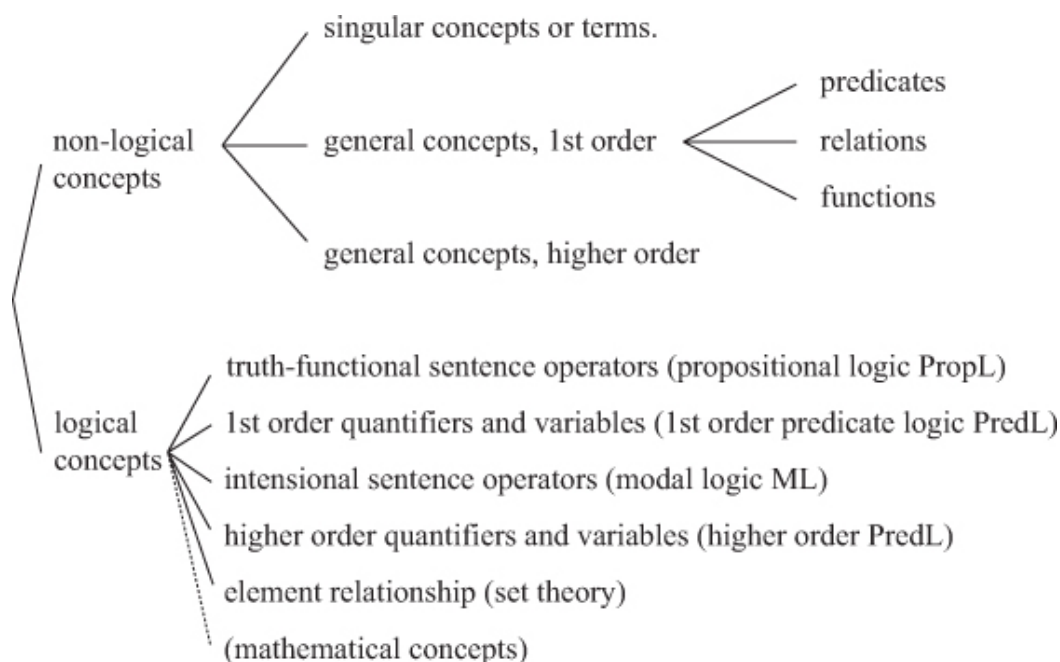


Figure 3.1-1 Division of concepts according to their logical type. [↩](#)

1. *Nonlogical concepts* denote or signify something in the world. By contrast, logical concepts have a purely *structural* function.

1.1. *Singular concepts* or *terms* always designate a single *individual* (or spatiotemporal location). Singular terms subsume (a) *proper names* (e.g. “Saul Kripke,” “Salzburg”), (b) *ostensive* or *indexical terms* (e.g. “I,” “now,” “here,” and “that person there”), and (c) *definite descriptions* (“that x which has characteristic property F”). *Formal notation*: atomic singular terms are called *individual constants* and are formalized by small letters $a, a_1, a_2, \dots, b, \dots$

1.2. *First-order general concepts* are applied to singular terms: this application is called *predication*, and through it, an *atomic sentence* is generated.

1.2.1. *First-order predicates* denote features which can be true of individuals or not. One may distinguish between *property concepts* such as “x is big,” and *kind concepts* such as “x is a person.” *Formal notation*: predicates are formalized using capital letters $F, G, \dots$ (or $F_1, F_2, \dots$). Predicates are *one-place*, that is, they have one argument *place*, which is written as a *variable* after the predicate letter: “Fx” for “x is an F.” For example, if “Fx” stands for “x is slim,” and “a” for “Peter,” then “Fa” stands for the sentence “Peter is slim.”

1.2.2. A (first-order) *n*-place *relation sign* denotes a relation between *n* given individuals. *Formal notation* by uppercase letters $R, Q, \dots$, for example, if “Rxy” stands for the two-placed relation “x is a brother of y,” “a” for “Peter,” and “b” for “Paul,” then “Rab” stands for the sentence “Peter is a brother of Paul.”

1.2.3. *Function signs*: a one-place (or *n*-place) function sign denotes a function which uniquely assigns to one (or *n*) individuals a further individual. *Formal notation* usually by lowercase letters $f, g, \dots$. For example, if “f(x)” stands for “the mother of x” and “a” for “Peter,” then “f(a)” is a singular term designating Peter’s mother. In mathematics, functions play a fundamental role as operations on numbers; for example, addition + is a two-place function, which assigns to the respective numbers *x* and *y* their sum $+(x,y)$, in infix notation written as “ $x+y$.” Moreover, all *quantitative* (e.g. physical) *properties* of individuals are functions: for example, $m(a)$ denotes the *mass* of the body *a*, and “ $m(a) = 5.6 \text{ kg}$ ” stands for “the mass of *a* is 5.6 kilograms.”

1.3. *Second (or higher)-order predicates and functions* are applied to sentences or predicates and thereby generate sentences. An example of a higher-order predicate is the concept of *truth* which applies to sentences.

2. *Logical concepts*: there are different kinds of logical concepts, and their principles are explicated in corresponding kinds of logics.

2.1. *Truth-functional sentence operators – Propositional logic PropL* (where $A, B, \dots$ stand for arbitrary propositions):

- the negation $\neg$ (read: $\neg A$ – *not A*),
- the conjunction $\wedge$ (read: $A \wedge B$ – *A and B*),
- the disjunction (the *inclusive* or) $\vee$ (read: $A \vee B$ – *A or B or both*),
- the so-called ‘material’ implication $\rightarrow$ (read: $A \rightarrow B$ – *if A, then B*),
- the ‘material’ equivalence $\leftrightarrow$ (read: $A \leftrightarrow B$ – *A if and only if (iff) B*),
- *convention*: “ $\leftrightarrow_{\text{def}}$ ” stands for “equivalent by definition”; similarly for “ $=_{\text{def}}$.”

Further sentence operators are thereby definable. For example, the exclusive *either-or* $\dot{\vee}$ is defined as $A \dot{\vee} B \leftrightarrow_{\text{def}} (A \vee B) \wedge \neg(A \wedge B)$ (*A or B, but not both*). The logical laws of these sentence operators are explicated by (non-modal) propositional logic.

A sentence operator, when applied to given sentences, generates a new, complex sentence. A sentence operator is called *truth functional* (or *extensional*) iff the truth value of the complex sentence is uniquely determined by the truth values of its component sentences (arguments). The truth-functionality of the sentence operators of PropL is expressed by their “truth tables”: every sentence *A* is either true or false; $\neg A$ is true iff *A* is false (else false); $A \wedge B$ is true iff both *A* and *B* are true (else false), etc. *Note*: the material implication $A \rightarrow B$ is defined as true iff *A* is false or *B* is true, that is, $A \rightarrow B \leftrightarrow \neg A \vee B$ is valid. The material implication is *weaker* than the natural language if-then, because its truth does not presuppose any connection between the contents of *A* and *B*.

2.2 *First-order quantifiers and variables – predicate logic PredL* : first-order *quantifiers* quantify over individuals with help of first-order *individual variables* $x, y, \dots (x_1, x_2, \dots)$. The two most important kinds of quantifiers are:

- the *universal quantifier* $\forall x$: read “ $\forall xFx$ ” as “For all x : Fx ,” and
- the *existential quantifier* $\exists x$: read “ $\exists xFx$ ” as “For at least one x : Fx .”

The class of individuals to which a quantifier refers is called its (semantic) domain D . So $\forall xFx$ says that all individuals in D have property F , and $\exists xFx$ that some individual in D has property F . When nothing further is stated, D is identified with the *universal* domain, that is, the class of all individuals in the given scientific area.

We say that the individual variable “ x ” is *bound* by the quantifier “ $\forall x$ ” or “ $\exists x$.” The subformula “ Fx ” in “ $\forall xFx$ ” is called the scope of the quantifier “ $\forall x$.” Formulas in which all individual variables are *bound* are called *closed formulas* or *sentences*. The expression “ Fx ” (for “ x is an F ”) is called an *open* formula, because the individual variable “ x ” is *free*, that is, not bound by any quantifier. Open formulas such as “ $Fx \wedge Gx$ ” express complex concepts.

A distinguished logical relation is *identity* “ $x = y$,” standing for “ x is identical with y ,” that is, “ x ” and “ y ” denote the same individual. The logic of sentences which are made up of singular terms, general first-order concepts, extensional sentence operators, and first-order quantifiers, is called first-order predicate logic PredL (introduction see [Klenk 1989](#) , [Copi et al. 2010](#)). First-order PredL contains PropL as a fragment and is one of the most important formal frameworks of philosophy of science, among other reasons because type-free set theory, upon which all of classical mathematics builds, is expressible in this framework.

2.3. *Intensional sentence operators* are not truth functional. An example is the necessity operator $\Box$ (“ $\Box A$ ” for “ A is necessary”): the truth value of “ $\Box A$ ” is not determined by the truth value of A . For instance, if A and B stand for the sentences “I have brown hair” and “ $2 + 2 = 4$,” respectively, then both A and B are true, but only $\Box B$ is true, while $\Box A$ is false, as the fact that I have brown hair is not necessary. Other examples are *deontic* sentence operators (“ A is obligatory”). The logic of intensional sentence operators is *generalized modal logic* (see [Hughes/Cresswell 1996](#)).

2.4. *Second (or higher) order quantifiers* quantify over states of affairs or properties and are dealt within second (or higher) order PredL. The logic of higher-order quantifiers is called logical *type theory*, which goes back to Russell ([1903](#)).

2.5. *Set-theoretical symbols*: sets are aggregations of individuals into a complex individual called a “set.” – *Formal notation*:

- “ $x \in y$ ” stands for “ x is an element of the set y .”
- $\{x: Fx\}$ stands for the set of all individuals x which possess the property F .
- $\{a_1, \dots, a_n\}$ stands for the set consisting of the individuals $a_1, \dots, a_n$.
- $(a_1, \dots, a_n)$ by contrast, stands for the ordered sequence consisting of $a_1, \dots, a_n$.

Note: sets are *invariant* with respect to (w.r.t.) permutations or repetitions of their elements, hence $\{a, b\} = \{a, a, b\} = \{b, a\} = \{b, a, a, b\}$ (etc.). By contrast, sequences or n -tuples consist in an ordering of individuals that allows for repetition: that is, $(a, b, b, c) \neq (a, b, c) \neq (b, a, c)$.

Set theory is a part of *informal* logic or metalogic, as “ $\in$ ” too is a logical concept in a broader sense. The axioms of type-free set theory (after *Zermelo* and *Fraenkel*) are expressed in the language of first-order PredL, but they go *beyond* purely logical principles (for an introduction, see van [Dalen et al. 1978](#) , [Hrbacek/Jech 1999](#)).

- “ $x \subset y$ (resp. $x \subseteq y$)” stands for “ x is a proper (resp. a proper or improper) subset of y .”
- “ $f: A \rightarrow B$ ” stands for “ f is a function which assigns to every individual x in A exactly one individual y in B ”.

Further set-theoretical notions are explained in the text.

2.6. *Mathematical concepts* , as with logical concepts, characterize *abstract-conceptual* objects or functions. They may be characterized either by internal axioms of mathematical theories or by set-theoretical definitions. Among the most important mathematical concepts are concepts for different kinds of numbers (e.g. $\mathbb{N}$ for the set of natural and $\mathbb{R}$ for the set of real numbers), and function concepts expressing operations on numbers (addition, multiplication, etc.).

3.1.2 Syntax and Semantics, Statements, and Models

Since Morris ([1946](#)), the dimensions of language are subdivided into *syntax* , *semantics* , and *pragmatics* . Syntax deals with the grammatical form of linguistic expressions, semantics with the meanings of expressions, and pragmatics finally, with those properties of expressions which are determined by their *context* . In this section, we outline some foundations of syntax and semantics.

From atomic concepts, one may construct terms and formulas of arbitrarily high complexity. The translation of a natural language sentence into a logical sentence is called *formalization*. Here are some examples:

(3.1-1) *Examples of formalizations:*

Translations: a – this animal, Rx – x is a raven, Bx – x is black.

- i. This animal is a raven and it is black. Formalization: $Ra \wedge Ba$.
- ii. All ravens are black. *Partial formalization:* For all x: if x is a raven, then x is black. *Formalization:* $\forall x(Rx \rightarrow Bx)$.
- iii. Some ravens are not black. *Partial formalization:* There exists an x which is a raven and is not black. *Formalization:* $\exists x(Rx \wedge \neg Bx)$.

Sentences with the same natural language syntax can have a very different logical syntax. In particular, the natural language copula “is” serves three different logical functions:

(3.1-2)

<i>Sentence</i>	<i>Formalization</i>	
	Ta	“is” as predication
i. Peter (a) is tall (T).	$a = f(b)$	“is” as identity
ii. Peter is the father (f) of Paul (b).	$\forall x(Rx \rightarrow Bx)$	“is” as inclusion of properties
iii. A raven (R) is a bird (B).		

In ([3.1-2](#))(iii), there is a *concealed* universal quantifier, which is revealed by formalization. The formalization of ([3.1-3](#)) reveals a *concealed* existential quantifier:

([3.1-3](#)) Peter (a) has a son. *Partial formalization:* There exists an x, such that x is the son (S) of Peter. *Formalization:* $\exists x Sxa$.

Logical formalization is an important aid to the advanced philosophy of science, to uncover the logical structure of sentences, and building on this, to discover logical connections. In this book, formalization is used rarely and as little as possible; for most purposes, partial formalizations such as in ([3.1-1](#)) above are sufficient.

In *semantics* , one differentiates between two kinds of semantic denotation: every expression has on the one hand a *meaning* , and on the other an object *reference* . (Terminologies differ, however; cf. [Kutschera 1976](#) , ch. 2; [Morris 2006](#) .) The meaning or intension of a concept is that which must be known to a speaker in order to *understand* the concept properly. The meaning of an expression *together* with the *facts* of the world determines the expression’s reference, that is, the real entity denoted by the expression (if there is no such entity, the reference is empty). There are controversial views on the question of wherein *exactly*

meaning exists (in the imagination, in thought, in the definition, or in usage?). The reference of a singular term is uncontroversially an individual. The reference of predicates is also controversial: some authors (e.g. [Carnap 1956b](#) , ch. I) identify the reference of a predicate with its *extension* , that is, the *set* of individuals which fall under the predicate; other authors (e.g. [Fodor 1990](#) , 93) identify the reference of a predicate with the *real property* or natural kind denoted by it.

The meaning of sentences is usually identified with the *proposition* expressed by the sentence. According to the traditional view, logically or analytically equivalent sentences express the *same* proposition. The reference of sentences is sometimes identified with truth values (e.g. in [Frege 1892](#)), and sometimes with events or real states of affairs. The following definition of “state of affairs” and “fact” is uncontroversial: every descriptive sentence expresses a *state of affairs* ; and if the sentence is true, the state of affairs really exists and is then a *fact* .

In *logical semantics* , which goes back to *Alfred Tarski* (1936a), one assigns to linguistic expressions formal object references, in the form of set-theoretical *interpretations* . Informal set theory functions in this case as a *metalanguage* , in which the semantic relationship between the *object language* and a hypothetical object structure is modeled. The concrete definition is as follows: an extensional *interpretation* of a logical language $\mathcal{L}$ is a pair (D,I), consisting of a domain of individuals D and an interpretation function I, which assigns to every individual constant a and variable x from $\mathcal{L}$ an individual I(a) in D, to every n-placed relation sign R in $\mathcal{L}$ a subset I(R) from D^n (where D^n is the set of all n-tuples of D-individuals), and to every n-placed function sign f in $\mathcal{L}$ a function $I(f):D^n \rightarrow D$. An extensional interpretation (D,I) makes an atomic formula Fa true, in short $(D,I) \models Fa$, iff $I(a) \in I(F)$. This basic truth definition is recursively extended to arbitrary formulas as usual (cf. [Machover 1996](#) , sec. 8.4; [Bergmann et al. 1998](#) , sec. 8.7).

The extensional interpretation of an object language can also be seen as a *relationship of translation* between a logical and a set-theoretical language – for example as follows:

(Prop. 3.1-1) *Translation (I) between logical and set-theoretical language* , where $I(a) = a$, $I(F) = F$, etc.:

Logical sentence	Set-theoretical sentence (translation)
Fa	$a \in F$
$\forall x(Fx \rightarrow Gx)$	$\forall x(x \in F \rightarrow x \in G)$, in short: $F \subseteq G$
$\psi(F)$	$F \in \psi$ (etc.)

In set-theoretical translations, then, individual constants and predicates or relation symbols are replaced by set terms which denote their extensions, and the element relationship $\in$ is inserted in between. This procedure is iterable for higher-order predications. Set-theoretical formalizations may be less intuitive than logical formalizations, but they have the advantage to be *type free* , that is, they allow for the translation of all sentences of (higher-order) type theory into the language of first-order PredL. For this reason, structuralist philosophers of science have argued that set-theoretical formalizations are preferable to logical formalizations (see [sec. 1.3.6](#)).

An extensional interpretation (D,I) of a language $\mathcal{L}$ is also called a *model* for $\mathcal{L}$. It plays the role of a “possible world” as expressible by means of $\mathcal{L}$. Often the models (D,I) of a sentence are written in the form $(D; a_1, \dots, a_n; R_1, \dots, R_m; f_1, \dots, f_k)$, in which instead of I, the designata of all individual constants, relational and function symbols assigned by I are listed. In this form, models are used by proponents of the (German) *structuralist* view and the (Anglo-Saxon) *semantic* view of theories. Here, a theory is represented by the set of its models, M(T) – a method that goes back to Suppes ([1999](#) , 249). By contrast, in the traditional “statement” view, a theory is represented by the set of its axioms, Ax(T), together with semantic characterizations of its concepts.

The (traditional) “statement view” and the (structuralist) “non-statement view” have been set in opposition by Stegmüller ([1976](#)). This “opposition” is largely unjustified, for two reasons. *First* , the statement view is not (as is sometimes claimed) only a purely syntactical, but also a semantical view,

inasmuch as the statements of $\mathcal{L}$ are taken as semantically interpreted propositions (cf. [French 2008](#) , 272). Since his turn in 1947, Carnap understood a linguistic framework always as a language $\mathcal{L}$ *together* with a semantic interpretation of $\mathcal{L}$ ([Carnap 1956b](#)). *Second* , model-theoretic theory reconstructions are not free from using statements (which is wrongly suggested by the name “non-statement view”). Rather, the class of models $M(T)$ of a theory T is defined by nothing but the set of its linguistic axioms $Ax(T)$, as follows: $M(T) =_{\text{def}} \{m \in M(\mathcal{L}) : m \models Ax(T)\}$, with $M(\mathcal{L})$ as the set of models for the language $\mathcal{L}$ in which T is formulated (cf. [Balzer et al. 1987](#) , ch. 1). Therefore, there exist straightforward *translation relations* between linguistic and structuralist theory-representations. Here is a simple example:

(Prop. 3.1-2) *Linguistic and model-theoretic theory representation* (simplest case):

Linguistically by T's axioms:	Model-theoretically by T's models:
Semantic ‘axioms’:	Models: $M(T)$ = the class of models
$A_1 : I(F) \subseteq D. A_2 : I(G) \subseteq D.$	(D,F,G) such that $(D,F,G) \models \{A_1, A_2, A_3\}$.
Factual axiom $A_3 : \forall x(Fx \rightarrow Gx).$	

Often translation relations are more complicated, since structuralist theories contain several ingredients; for details see Andreas ([2010](#)) and Schurz ([2014b](#)).

3.1.3 The Content Type of Concepts

This classification is particularly significant in the philosophy of science. It applies only to *nonlogical concepts* . [Fig. 3.1-2](#) shows an *overview* .

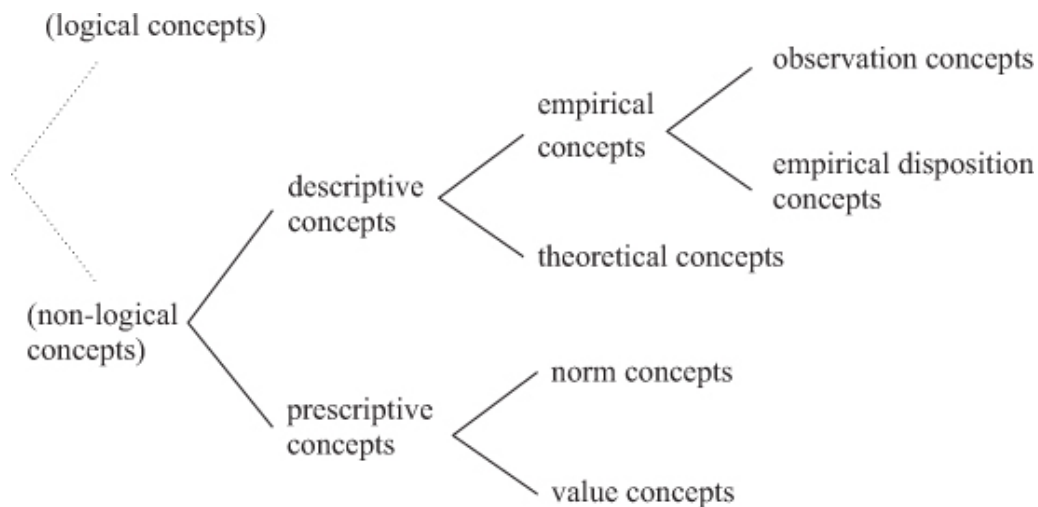


Figure 3.1-2 Classification of concepts according to their content type. ↵

First, we have to distinguish between descriptive and prescriptive concepts. *Prescriptive* concepts consist of norm and value concepts ([Hare 1952](#)). There are only a few prescriptive concepts: the *normative* concepts of something being obligatory, forbidden, or permitted, and the qualitative, comparative, and quantitative *value* concepts (X is ethically or aesthetically valuable, or more valuable than Y , or valuable to a certain degree). Philosophy of science deals primarily with *descriptive* concepts and sentences which describe the facts of the world. There is an enormous number of (singular and general) descriptive concepts of various types.

The *most important* distinction within the descriptive concepts is the one between empirical and theoretical concepts, and within the empirical concepts the one between observation concepts and empirical dispositional concepts.

(1.) A general *observation concept in the narrow sense* (i.n.s.) expresses a property that is perceptible to the senses, or a complex of such properties. As we saw in [sec. 2.9.2](#), a property perceptible to the senses is one which can be *ostensively learned* by humans, independently of their background knowledge. *Examples*: “x is red,” “x is a raven,” “x is longer than y,” “the needle of the measuring instrument x is pointing to 5.” The classical and the early logical empiricists demanded that all descriptive concepts be observational concepts. In fact, only a *few* concepts are genuine observation concepts in this narrow sense.

Note the following subtlety: strictly speaking, a property as such is not perceptible, but only a certain *instantiation* of a property (a so-called *trope*) in a perceptible individual. “Being red,” for example, expresses an observational property as applied to visible objects; but in application to non-observable individuals such as microbes or molecules, the property of being red is not observable, as these individuals are too small to be visible. For this reason, Quine was right when he stressed that the primary level of “observability” is not that of concepts, but of “occasion sentences” like “this here is red” (1960, 35ff; 1992, 3). Putnam ([1962](#), 242f) saw in this fact a fundamental obstacle to a clear definition of observation concepts. Our criterion of ostensive learnability, however, solves this problem as follows: a primitive property is ostensively learnable, and thus observable i.n.s., iff it possesses at least *some* perceptible instantiations in perceptible individuals which may serve as training instances.

(2.) An *empirical dispositional concept* expresses an empirical disposition. By this, we understand the *lawlike regularity* of an individual to produce, under certain empirical circumstances or trigger conditions, a certain observable *result* that is also called the disposition’s “manifestation” ([Schrenk 2017](#), 45). Example: the empirical dispositional concept “x is soluble in water” is by definition equivalent to the following lawlike hypothesis about “x” which generalizes over time points: “whenever x is or were placed in water, it will or would dissolve.” Lawlikeness and counterfactual conditionals (“would”) go beyond the observable and presuppose inductive projectibility (see [sec. 3.6](#), 6.5.3). Nevertheless, dispositional concepts are *empirically operationalized* concepts, whose meaning is not dependent on a theory. Further examples: “x is hard,” “x weighs 4 kg on a pair of scales,” “person x is a German speaker.” Empirical dispositional concepts correspond to what the *operationalists* (e.g. [Bridgeman 1936](#)) demanded of *all* scientific concepts. In fact, though, theoretical concepts are not reducible to empirical dispositional concepts (see [sec. 3.10](#)).

(3.) An *empirical concept i.n.s.* is an observation concept or an empirical dispositional concept.

(4.) A *theoretical concept in the wide sense* (i.w.s.) is any descriptive concept which is not an empirical concept i.n.s., but still has a meaning. A theoretical concept i.w.s. does not have to be part of a scientific theory; it can also be based on speculation (e.g. “the prime cause of all being”).

(5.) A *theoretical concept i.n.s.* is a theoretical concept i.w.s., whose meaning is *introduced* by a scientific theory. If we want to make explicit the reference to a specific theory T, we speak of a *T-theoretical concept* ([Sneed 1971](#); refinements in [sec. 5.3.2](#)). *Examples*: physical force, kinds of molecules, biological genotype, psychological character, intention, and social order.

Convention: unless otherwise stated, we understand empirical concepts and theoretical concepts respectively in the narrow sense (i.n.s.).

(6.) As discussed in [sec. 2.9.1](#) (pro-5), a number of philosophers of science use the notion of an empirical concept in the wide sense (i.w.s.), meaning any concept that expresses a *measurable property* or ‘datum.’ Empirical concepts i.w.s. are theory dependent – their meaning depends on theories which describe the functioning of the measuring instruments and are presupposed to be unproblematic. The border between empirical concepts i.w.s. and theoretical concepts i.n.s. can only be drawn *pragmatically* and in a *context-dependent* way. By contrast, there exists an *objective* but gradual border between empirical concepts i.n.s. and theoretical concepts i.w.s.

3.1.4 The Gradation (Scale) Type of Concepts

Most properties in our world do not only occur or not occur; they come in certain *degrees*. The gradation or scale type of a concept reflects how finely the concept is graduated. Accordingly, scientific concepts may be classified as in the *overview* in [fig. 3.1-3](#).

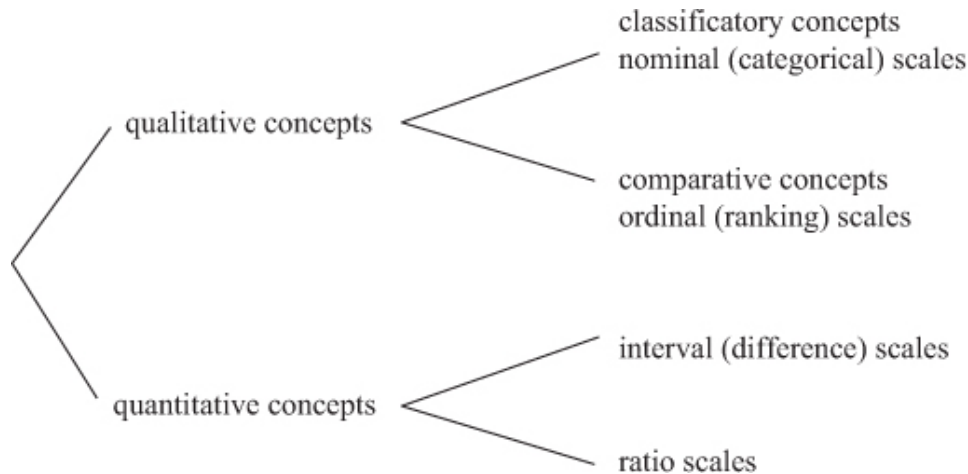


Figure 3.1-3 Classification of concepts according to their gradation type. [↗](#)

Qualitative concepts are not graduated, or more precisely, they are graduated only in the *yes-no* sense; comparative concepts are graduated in the sense of an *ordering*, and quantitative concepts are the most strongly graduated, that is, in a *continuous* way. In science, the graduation of a property is measured in terms of numbers, and we speak of the scale type of concepts (e.g. [Krantz et al. 2006](#), ch. 1; [Bortz 1985](#), 28–33). The *finer* graduated a concept is, the *sharper* the information we express with it, and the *stricter* the tests to which the hypotheses using it can be subjected. The types of gradation will now be discussed.

3.1.4.1 Qualitative-Classificatory Concepts and Nominal Scales

With the help of qualitative or classificatory concepts, we divide a domain D of objects into certain subgroups $D_1, \dots, D_n$, which are defined by their having certain properties or belonging to certain categories $C_1, \dots, C_n$. Here, it is demanded that: (a) the classification is *disjoint*, that is, the categories C_i must be mutually exclusive (the subgroups must not overlap), and (b) the classification is *exhaustive*, that is, the union of the subgroups must contain all of D .

The set $F =_{\text{def}} \{C_1, \dots, C_n\}$ of the categories of a classification is called a (classificatory) *concept family* ([Carnap 1971](#), 43ff). In the simplest case, an analytically independent concept Fx forms a two-valued or *binary* concept family $\{Fx, \neg Fx\}$: for example, x is either human or not. The conditions of disjunctivity and exhaustivity hold in this case purely logically. There are also concept families with more than two categories; for example, the family of color concepts $\{\text{red, yellow, green, and blue, } \dots\}$, or the family of social classes $\{\text{lower class, middle class, and upper class}\}$. The conditions of disjunctivity and exhaustivity are in these cases based on analytical conventions.

Every concept family F corresponds to a family concept or *attribute* X_F . For example, the concept family $\{\text{lower, middle, and upper class}\}$ corresponds to the attribute “social class of x .” Formally, X_F is a higher-order *function*, which assigns to every object in D its characteristic category, that is, $X_F : D \rightarrow \{C_1, \dots, C_n\}$. If we replace the categories C_i by *natural numbers* i , for example, “lower, middle, upper class” by “1,2,3,” then we obtain a nominal (or categorical) *scale*. For nominal scales, however, the use of numbering is inessential and arbitrary. In mathematics, a function X which when applied to various objects can take on various *numerical values* is also called a *variable*. Simple binary properties (F) correspond to a *binary* variable X_F with values F and $\neg F$. Observe that mathematical variables are functions; they should not be confused with variables in the logical sense.

3.1.4.2 Comparative concepts and ordinal scales

With the help of comparative concepts, the objects of a domain D are *ordered* according to the grade of a given property, that is, they are divided into *ranking groups* $D_1, \dots, D_n$ (see [fig. 3.1-4](#)): D_1 contains all D -objects with the highest grade of the given property, D_2 all D -objects with the next highest grade, etc. The

objects (“o”) within a rank group have the same grade. If the rank groups D_i are assigned natural numbers i , we have an *ordinal* or *ranking scale*. The corresponding assignment function $r: D \rightarrow \{1, \dots, n\}$ is an ordinally scaled variable, also called a ranking function.

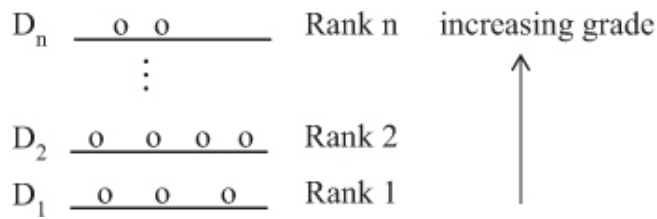


Figure 3.1-4 Ordinal scale of D . [↗](#)

The ranks of an ordinal scale give us only comparative information; quantitative differences between the ordinal ranks are unknown or undefined. In the *Mohs scale*, for example, the hardness of minerals is determined by a reciprocal *scratch comparison*: if mineral A scratches mineral B, but not vice versa, then A has a higher degree of hardness. Mohs used this method to devise a ten-point hardness scale: 1 (talcum), 2 (gypsum), 3 (calcium carbonate), up to 9 (corundum), and 10 (diamond). We must not conclude from this that the difference in hardness between talcum and gypsum is for example one-ninth of the difference between talcum and diamond. Similarly, the intensity of *psychological* attitudes can typically be measured only by an ordinal scale: for example, I can find one dish *tastier* than another without being able to say how much *tastier* it is.

An important method used in the empirical social sciences for *attitude surveys* consists in asking test persons (tps) to place their opinions on a given topic into a *preset* ordinal scale. A questionnaire on the attitudes of tps toward old age could for example contain the following question:

Please tick where appropriate: I believe that securing an old age pension is:

- ☐ very important ☐ quite important ☐ don't know
☐ not so important ☐ not important at all

The attitudes of the persons surveyed are thereby measured by an ordinal scale with five ranking levels, from not important at all (1) up to very important (5).

The basic logical concept of ordinal scales is that of a *smaller-or-equal* relation $\leq_P$ in regard to the given property P : $x \leq_P y$ states that the P -grade of x is less than or equal to that of y . With help of “ $\leq_P$,” one defines the relation of *P-equality* as $x \equiv_P y \leftrightarrow_{\text{def}} x \leq_P y \wedge y \leq_P x$ (x 's P -grade equals y 's), and the *strictly smaller* relation as $x <_P y \leftrightarrow_{\text{def}} x \leq_P y \wedge \neg y \leq_P x$ (x 's P -grade is smaller than y 's). The concepts $\leq_P$, $\equiv_P$, and $<_P$ are operationalized by *empirical* operations of comparison. In order to generate a *unique* ranking scale, these comparison operations must fulfill at least the axioms of a total *quasi-order*: *reflexivity* ($\forall x: x \leq_P x$), *transitivity* ($\forall x, y, z: x \leq_P y \wedge y \leq_P z \rightarrow x \leq_P z$), and *connexivity* ($\forall x, y: x \leq_P y \vee y \leq_P x$), that is, everything is comparable to everything else. Axioms and definitions entail the *anti-symmetry* of $\leq_P$ ($\forall x, y: x \leq_P y \wedge \neg x \equiv_P y \rightarrow \neg y \leq_P x$) and the fact that $\equiv_P$ is an *equivalence relation*, that is, fulfills the axioms of reflexivity, symmetry, and transitivity. (Note: an *order* in the mathematical sense arises from a quasi-order if no two D -objects occupy the same rank.)

The axioms of transitivity and connexivity are prerequisites for ordinal scaling. Whether they are at least nearly fulfilled is an *empirical* question (only reflexivity and anti-symmetry are fulfilled analytically). It is well known that the empirical preferences of tps are not always transitive (cf. [Garnham/Oakhill 1994](#), 184ff). In addition, all empirical *pair comparisons* have a certain *inaccuracy*. The relation of *approximate equality*, though, is no longer transitive when the grade differences between objects are very small. As an example, assume that a is 0.6 mm taller than b , b is 0.6 mm taller than c , but that only height differences of

more than 1 mm are observable. Then we observe that a is approximately equal to b, and b to c, but a is taller than c. If we place objects of this sort in a *preset* ordinal scale, we make their rank relations artificially transitive: the result may then depend on the order in which the objects are considered by the tp.

3.1.4.3 . Quantitative concepts and scales

These express the grade of a property P by means of a *numerical value* , along with a *unit of measurement* . Logically speaking, a quantitative concept is a function $f(x)$, which assigns to every object x in domain D a numeric value $r \in \text{Num}$ along with a unit of measurement e_f ; for example, $\text{mass}(x) = 5 \text{ kg}$, or $\text{temperature}(x) = 20^\circ \text{ Celsius}$. If $f(x)$ is a *discrete* magnitude, then f assigns integers to the objects; if $f(x)$ is a *continuous* magnitude, then f assigns real numbers to the objects. Almost all physical magnitudes are continuous; exceptions occur in quantum mechanics. Typical discrete magnitudes are counting numbers, for example, population sizes. In metaphysics, the nonspecified magnitude (e.g. “mass”) is also referred to as a determinable, and its numerical instantiation (e.g. “mass of 3.791 kg”) as a determinate property ([Wilson 2023](#)).

Directly observable are at best *comparative* magnitude relations: if we determine the length of a given distance by using a measuring tape, we can see whether this distance is longer, shorter, or approximately equal to certain scale divisions on the measuring tape. Converting a comparative concept into a quantitative concept by introducing an (empirical) measuring operation is called its *metrization* . In order to make a comparative property quantitatively measurable, certain empirical conditions must be satisfied, which are explicated axiomatically in *metrization theory* (see [Krantz et al. 2006](#)). On the other hand, any empirical measurement operation is based on certain *arbitrary conventions* , to which the results of respective measurements are relativized. Interval and ratio scales differ from each other in the amount of these arbitrary conventions. In an *interval scale* , both the *zero point* of the scale and the *unit of measurement* are set by convention. In a ratio scale, on the other hand, only the unit of measurement is conventional, while the zero point is empirically determined.

The empirical temperature scale, for example, is an interval scale. It is based on the fact that certain liquids, for example, mercury in a glass column, expand uniformly as the temperature increases. But there is no zero point of warmth or coldness which might be discerned through mere observation. In the *Celsius scale* ($^\circ \text{C}$), the freezing point of water is used as a conventional zero point, and the scale unit of 1°C is conventionally defined as one 100th of the length of expansion of mercury in a glass column as it rises from the freezing point to the boiling point of water. In the *Fahrenheit-scale* ($^\circ \text{F}$), the freezing point of water is set at 32°F , and the scale unit of 1°F corresponds to $5/9^\circ \text{C}$. So physically, it makes no sense to say that a liquid at 20°C is *twice as warm* as one at 10°C . This relationship depends on the artificial choice of the zero point; in the $^\circ \text{F}$ scale, the same relationship comes out as 68°F to 50°F . If, however, we have three liquids; a, b, and c at 10°C , 20°C , and 30°C , respectively, then it certainly does make sense in physical terms to say that the *temperature difference* between b and a is exactly as great as the one between c and b, and half as great as the one between a and c, as this is valid *independently* of the choice of zero point: in the difference between two temperature values, the zero point is canceled out. Because only statements about ratios of temperature intervals are objective, while statements about ratios of temperatures lack objective content, we talk of interval scales. Measurements of *position* and *time* also use interval scales, because the zero points of spatial and temporal coordinate systems are set by pure convention. We cannot say for example “the year 2000 is twice as late as the year 1000,” as this is relative to the birth of Christ as the starting point of the calendar. If our calendar began in the Middle Stone Age, we would have instead the years 12000 and 11000.

In *ratio scales* , unlike interval scales, the absolute zero point is empirically determined. For example, mass, volume, length (as opposed to spatial position), and duration (as opposed to temporal position) are ratio-scaled magnitudes. Physically, it makes sense to say that Peter, at 100kg, weighs twice as much as his 50 kg little brother. Because such ratio statements have objective content, we speak of ratio scales. Similarly, quantitative magnitudes in economics, such as amounts of money or goods, are based on ratio scales.

These magnitudes are also called *extensive* magnitudes, because when joining or “concatenating” two objects (a, b) into a whole (a o b), the resulting magnitude grows monotonically, that is, it holds that $f(a \text{ o } b) > f(a), f(b)$. ¹ Magnitudes such as temperature, for which this is not the case, are by contrast called *intensive* magnitudes – the temperature of say, water, is not increased if we increase the amount of water.

The metrization of extensive magnitudes leads to ratio scales. What is still arbitrary in ratio scales is the (conventional) choice of a *unit*. It makes no sense to say that the numerical value of Peter's weight is 100 – this is only valid if the unit of mass is assumed to be the kilogram; if we chose the gram as unit instead, the numerical value of Peter's mass would be 100,000. One subtle problem is the justification of the assumption that the chosen unit remains *constant over time* (see on this [Carnap 1966](#), ch. 8–9).

Only in simple *counting scales* is the unit naturally determined as one individual or item; such counting scales are also called *absolute scales*. The difference between various types of scales is mathematically described in terms of *admissible scale transformations*. The coarser a scale, the more scale transformations are admissible; for details see Krantz et al. ([2006](#), 11).

In theories of metrization, one proves, for example, the following: an extensive comparative magnitude relation $\leq_p$, for example, “is longer than,” over a domain of objects D is metrizable in the form of a ratio scale iff $\leq_p$ is a monotone quasi-order over D which satisfies the so-called *Archimedean* condition (see [Krantz et al. 2006](#), 73). The latter condition demands that every object b, no matter how large, can be balanced out by a sufficient number of copies of an arbitrary object a, no matter how small a is. For social preference relations, the Archimedean condition is often not satisfied – for example, one would be reluctant to say that the value of a human life might be balanced out by a sufficient number of banknotes (cf. [Schurz 1997](#), sec. 11.5.2).

Strictly speaking, the results of attitude measurements (e.g. 5 = very important, 3 = don't know, and 1 = unimportant) only have the level of an ordinal scale, or they lie somewhere “between” ordinal- and interval-scale levels. In spite of this, they are often evaluated using statistical tests which presuppose interval scales. To assess whether such an idealization is harmless in a given case, it is necessary to reevaluate the results of interval scale-based tests by an ordinal scale-based (or “non-parametric”) test (see [Bortz 1985](#), 31f, 178f).

So far, we have been concerned with *fundamental empirical* metrization. One speaks of *derived* metrization when the metrization of a concept is traced back to the metrization of other concepts, for example, density to the quotient of mass divided by volume. One speaks, finally, of *theory-dependent* metrization, when the metrization is based on a certain theory, for example, from physics ([Balzer 1985b](#), 6). In theory-dependent metrizations, a change in the scale level of an initially empirically defined metric may occur for theoretical reasons. An example of this is the *absolute* temperature scale after Kelvin ($^{\circ}\text{K}$), which postulates the absolute zero point of temperature at $-273.15^{\circ}\text{C} = 0^{\circ}\text{K}$. This assumption resulted from the *kinetic theory of gases*, according to which heat and temperature are to be attributed to the movement of molecules. In ideal gases, the volume (V) increases in a *linear* way with rising temperature ($V = k \cdot T$). If one extrapolates this relationship in a T×V-coordinate system to the left into the region of negative $^{\circ}\text{C}$ -temperatures, one reaches a temperature of -273.15°C , at which the volume of an ideal gas would have to be *zero*: here, the gas molecules would be crammed into an infinitely small space, and thus would be unable to move. Kelvin therefore identified this point as the absolute zero of temperature. His hypothesis that no cooling process could go below this absolute zero point has since been confirmed many times.

3.2 Classification of Sentences According to Their Content Type

Sentences may also be classified into several different dimensions. Most significant in philosophy of science is the classification of sentences according to their content type, which is treated of here, and according to their degree of generality, which is dealt with in sec 3.5. [Fig. 3.2-1](#) gives an *overview*.

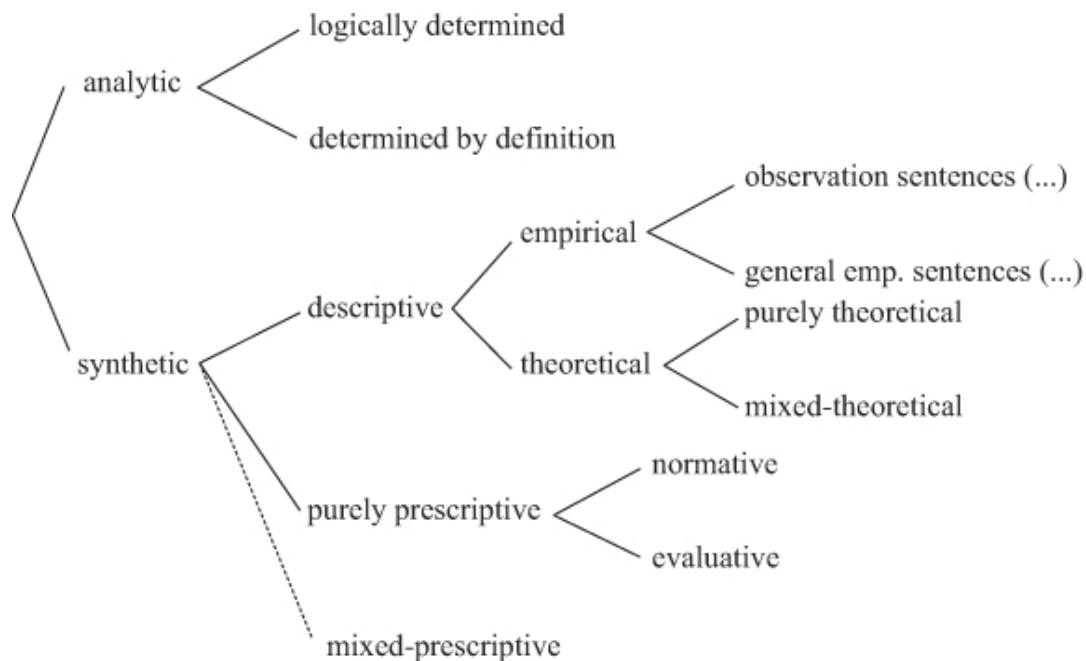


Figure 3.2-1 Classification of sentences according to their content type. [↗](#)

Analytically true sentences are those whose truth is determined by the laws of logic or by extralogical conventions of linguistic meaning, independently of the factual constitution of the world. Analytically false sentences are, by analogy, those whose falsity is determined in this way, and *analytic sentences* are those whose truth value (true or false) is determined in this way. Analytic sentences may be broken down further into logically determined sentences and sentences determined by definition or extralogical meaning convention ([sec. 3.3](#) – [3.4](#)).

By contrast, *synthetic* sentences are those sentences whose truth value is not analytically determined, but is dependent on the factual constitution of the world, or in the case of prescriptive-synthetic sentences, on our moral interpretation thereof. As with our content-based classification of concepts, we first divide synthetic sentences into descriptive and prescriptive ones; more precisely, into purely descriptive sentences, purely prescriptive sentences, and mixed sentences with descriptive and prescriptive components. Descriptive sentences tell us something about the factual composition of the world; prescriptive sentences on the other hand express propositions about norms or values. In the following subsections, we refine these distinctions.

3.2.1 Descriptive versus Prescriptive Sentences

Not every sentence containing a prescriptive concept is thereby prescriptive. Indeed, whenever a prescriptive concept lies within the *scope* of a *subjective attitude operator* – that is, a sentence operator expressing a subjective attitude toward a state of affairs – the prescriptive concept loses its prescriptive force. So for example, the sentences

Peter believes that stealing is wrong, – or
 Peter believes that you shouldn't steal,

are empirically descriptive, as they state something about Peter's factual attitude to values or norms. The same goes for all subjective attitude operators, like for example: person X believes, thinks, feels, says, wishes,..., that A. The significance of this fact for the value-neutrality thesis was worked out in [sec. 2.5.1](#) . We define:

(Def. 3.2-1) (a) A sentence is *descriptive* iff every prescriptive symbol occurring in it lies within the scope of a subjective attitude operator.

(b) A sentence is *purely prescriptive* iff all of its descriptive components lie within the scope of a prescriptive operator.

(c) A sentence which is neither descriptive nor purely prescriptive is called *mixed-prescriptive*. These sentences contain both descriptive components outside the scope of any prescriptive operator and purely prescriptive components outside the scope of any subjective attitude operator.

(3.2-1) *Examples of descriptive sentences* (logical structure at the right; prescriptive operators underlined):

- i. Peter steals. – Steals(Peter).
- ii. Peter believes that stealing is forbidden – Believes(Peter, Forbidden (stealing)).

Examples of purely prescriptive sentences:

- iii Stealing is forbidden. – Forbidden (stealing).
- iv If an action is under all circumstances morally bad, then it is forbidden. – $\forall a,c$ (Bad (a,c) $\rightarrow$ Forbidden (a)).

Examples of mixed-prescriptive sentences:

- v For persons suffering from hunger, stealing is allowed. – $\forall x$ (Suffering-hunger(x) $\rightarrow$ Permitted (Steals(x))).

Two remarks: (a) implicational connections between purely prescriptive sentences as in iv are themselves purely prescriptive. (b) Def. 3.2-1 is a *syntactic* definition; it can be refined by the method of *essential* sentence types in [def. 3.5-1](#).

3.2.2 Observation and Basic Sentences, Empirical and Theoretical Sentences

Building on the content-based classification of concepts in [sec. 3.1.3](#), we classify descriptive sentences according to their content type. Here, we have to address two subtleties which are rarely discussed in the literature. They have to do with the fact mentioned in [sec. 3.1.3](#), that instantiations of observable properties are only observable if the individuals possessing them are also observable.

First subtlety: an *empirical* sentence is usually defined as a sentence which (apart from logical concepts) contains only empirical concepts (cf. [Tuomela 1973](#), 23f; [Kutschera 1972](#), 260, 292). But this is insufficient. Consider a universal sentence “All Fs are Gs,” which contains only the observation concepts F and G. This universal sentence can only be called empirical if it refers solely to perceptible individuals. Whether this is the case depends on the concept F: only if it follows from the meaning of F that F can be applied *only* to perceptible individuals is the universal sentence an empirical sentence. For example, “all ravens are black” is a universal empirical sentence, because it follows from the meaning of “raven” that only perceptible individuals can be ravens. However, “all loose objects on Earth fall downwards” is not a universal empirical sentence if by “objects” we are also talking about microscopically small particles. The same applies to existential sentences. Lewis ([1970](#), 79) and van Fraassen ([1980](#), 13f) concluded from this problem that one cannot ascertain the empirical versus theoretical nature of a sentence from the meaning of its concepts. But as we will show now, this is indeed possible. For this purpose, we introduce the following *definition*: a quantifier in a given sentence has an *empirical range* iff this sentence is analytically equivalent with its quantificational restriction to observable individuals only. For a universal sentence $\forall xA$, this restriction is defined as the sentence $\forall x(Ox \rightarrow A)$, and for an existential sentence $\exists xA$ as $\exists x(Ox \wedge A)$, where “ Ox ” stands for “ x is observable.” Otherwise, the quantifier has a *theoretical range*.

Second subtlety: observation sentences talk only of a finite or manageable number of individuals, because what humans can observe is of finite and limited complexity. It would be inadequate to limit observation sentences to *singular sentences* – that is, sentences *without* quantifiers – which contain only observation concepts. *Localized existential sentences* of the form “at point k in space-time, there is such-and-such an observable thing” are also observation sentences; for example, “in this basket (now) there is a red apple” ([Popper 1935](#) /2002, sec. V.28, speaks of “singular existential sentences”). This is the case because the

existence quantifier in such a sentence has empirical range, and because the limited segment of space-time is small enough to afford us a visual overview. For similar reasons, even *localized universal sentences* have to be counted as observation sentences, which have the following form: all (observable) individuals in this limited area of space-time have this-or-that observable property”; for example, “All apples in this basket are red.” Once again, the universal quantifier must have an empirical range; in addition, it is important that the individuals satisfying the antecedent predicate have a perceptible spatial extension that does *not overlap* with other individuals. Given these conditions, it follows from the meaning of “apple” and “basket” that there is only room for a finite and manageable number of apples in the described basket, so that we can determine the truth value of the sentence by observation.

This brings us to the following definition:

(Def. 3.2-2) If S is a synthetic-descriptive sentence, the following holds:

- a. S is an *observation sentence* iff apart from logical concepts, S contains only observation concepts, and S is either a singular sentence, or a *localized* existential or universal sentence, whose quantifiers possess empirical range.
- b. S is an *empirical sentence* (i.n.s.) iff apart from logical-mathematical concepts, S contains only empirical concepts i.n.s., and its quantifiers possess empirical range.
- c. S is a *theoretical sentence* i.w.s. iff S contains theoretical concepts i.w.s. or its quantifiers possess theoretical range.
- d. S is a *T-theoretical* sentence (or theoretical i.n.s.) iff S is a theoretical sentence i.w.s. and the theoretical concepts of S are introduced by the scientific theory T (refinements in [sec. 5.3.2](#)).

In analogy to [sec. 3.1.3](#) , we speak of an empirical sentence i.w.s. if apart from logical-mathematical concepts, it contains only empirical concepts i.w.s. (and quantifiers with an empirical range i.w.s.). Empirical sentences i.w.s. are dependent on unproblematic pre-theories and reflect what scientists pragmatically call ‘data.’ *Convention:* unless otherwise stated, we understand observation, empirical, and theoretical sentences i.n.s.

The class of empirical sentences is much more *comprehensive* than the class of observation sentences. Apart from observation sentences, it also comprises: (i) *universal* empirical sentences like for instance “all ravens are black,” (ii) singular empirical *disposition* sentences like “this is soluble in water,” and (iii) general empirical disposition sentences like “all sugar is soluble in water.” An important subgroup of empirical disposition sentences are empirical *measurement* sentences like “this tree is (rounded off) 10.52 meters high,” or “fully grown fir trees are more than 10 meters high.”

Theoretical hypotheses need not always be universal sentences. Singular or existence sentences can also be theoretical. For example, the singular sentence “this mineral is radioactive” expresses a theoretical hypothesis which can only be tested by means of theory-led measurements using complicated instruments. Finally, *theories* are *systems* of theoretical sentences (ch. 5). We stress here that *theories* or theoretical sentences may *also* contain empirical concepts, and even *must* do so in order to have empirical content, because some authors (e.g. [Putnam 1962](#) ; [Chalmers 1986](#) , 36; [Quine 1992](#) , 6–8) give the misleading impression that a concept, once it occurs within a theory, becomes a theoretical concept. According to our definitions, this is clearly not the case (recall [sec. 2.9.1](#) , contra-7).

Narrower still than the category of observation sentences is the category of *basic sentences* : basic sentences in the sense of Carnap ([1950b](#) , 67) and Hempel ([1965](#) , 277) are empirical atomic sentences or their negations, for example, this tree is (not) green. Basic sentences in the sense of Popper ([1935](#) /2002, ch. 5, sec. 28) are localized existential sentences, whose matrix is a conjunction of basic formulas, for example, “there is a black raven there.” All basic sentences are observation sentences, but not vice versa. In the Vienna Circle, it was hoped that basic sentences would serve to demarcate the *immediately* observable from that which can only be observed in a *mediated* way, via deduction from that which is immediately observable. For example, disjunctions such as “there is a crow or a raven sitting there” are not immediately observable, but they can be deduced from immediate observation sentences. Closer analysis, however, shows that this demarcation is problematic.

3.3 Logical Truth and Deductive Logic

According to philosophical tradition, logical truths are understood to mean “necessities of thought.” Both the history of philosophy and cognitive psychology show, however, that our *intuitive* ideas about necessities of thought are unreliable. Modern logic replaces the intuitive concept of necessity with exact concepts. The distinction introduced in [sec. 3.1.1](#) between logical and nonlogical concepts is of fundamental importance for the modern definition of logical truth. We will now present a general characterization of logical truth, which is applicable to any system of logic.

([3.3-1](#)) *Logical truth – preliminary characterization*: a sentence is logically true iff its truth follows logically from its syntactic structure and the meaning of its logical concepts alone.

In an equivalent reformulation of this idea, one can say: a sentence is logically true iff its truth follows from its *logical form* (this characterization goes back to [Wittgenstein’s Tractatus of 1921](#)). We find the logical form of a sentence by replacing all of its nonlogical concepts with metalanguage *variables* . The logical form of a sentence is what we get after the *formalization* of the sentence ([sec. 3.1.2](#)), but without any natural language interpretation of its nonlogical terms.

The characterization ([3.3-1](#)) is only preliminary, as the words “follows logically” in the defining phrase make it circular. The essential thing about a logical sentence form is that its nonlogical symbols may be *interpreted* in arbitrary ways. If every such interpretation makes the sentence true, then the sentence is logically true. We can therefore conclusively define the logical truth of a sentence, building on the idea in ([3.3-1](#)), as follows:

(Def. 3.3-1) A sentence is *logically true* (for short: *L-true*) iff its (logical) sentence form is true for every possible interpretation of its nonlogical symbols. We then also call this sentence form logically true.

Definition ([3.3-1](#)) goes back to Tarski ([1936b](#)). It is easily applicable to specific systems of logic, by specifying the permissible interpretations. In PropL, these are the assignments of truth values to atomic sentences. In PredL, these are the interpretations (D,I) explained in [sec. 3.1.2](#) . Note that in possible extensional interpretations (D,I), not only the interpretation function $I(\phi)$ for nonlogical symbols ϕ may vary, but also the individual domain D. Here are some examples:

(3.3-2) *Example of an L-true sentence*:

If everybody is mortal, then nobody is immortal.

Sentence form: if all Fs are Gs, then there exists no F which is not a G.

Formalization: $\forall x(Fx \rightarrow Gx) \rightarrow \neg \exists x(Fx \wedge \neg Gx)$

Further examples: $A \vee \neg A, A \rightarrow (A \vee B), \forall x Fx \rightarrow Fa, Fa \rightarrow \exists x Fx$.

The sentence form ([3.3-2](#)) is L-true, because it is true for every (extensional) interpretation of “F” and “G.” In other words, it is true in every possible world, as extensional interpretations correspond to possible worlds.

In analogy to [def. 3.3-1](#) , one defines a sentence as *logically false* (L-false) iff its sentence form is false for every possible interpretation of its nonlogical symbols. Examples of L-false sentence forms are $p \wedge \neg p$ (e.g. it will rain today and it will not rain), or $p \leftrightarrow \neg p$.

L-false sentences are also called *inconsistent* , or *contradictory* , because one may derive contradictions from them of the form “p and not p.” Sentences which are not L-false are called *consistent* . Consistency is also defined for *sets* of sentences: a sentence set is called consistent iff at least one interpretation verifies all sentences in the set; otherwise, it is called inconsistent (or contradictory). Finally, a sentence which is neither L-true nor L-false is called logically *contingent* . Examples of logically contingent sentences are Fa (e.g. Peter is ill) or $\forall x(Fx \rightarrow Gx)$ (e.g. all ravens are black). Most sentences are contingent.

With the help of the same method, we define the logical *validity* of arguments or inferences, which consists in the fact that the *truth* is transferred from the premises to the conclusion with *certainty* , that is, in all possible interpretations or ‘worlds’:

(Def. 3.3-2) An inference is *logically valid* (for short: valid) iff for the logical *form* of this inference the following holds: every possible interpretation which makes all premises true, also makes the conclusion true. – *Equivalently expressed*: ... iff there is no interpretation which makes all the premises true, but the conclusion false. – If an inference is valid, then we say that its conclusion *follows* (logically) from its premises.

(3.3-3) *Example of a logically valid inference*: Inference form:

Premise 1: All humans are mortal. $\forall x(Hx \rightarrow Mx)$

Premise 2: You are a human. Ha

Conclusion: Therefore, you are mortal. Ma

As with logical truth, logical validity is a matter of *logical form*. This fact implies that the validity of an inference does *not* depend on the *truth* of its premises (a common misunderstanding). The fact that an inference is valid only means: *if* its premises *were true*, then its conclusion *would* also have to be true. An inference with false premises can also be valid, like for example: “All humans are rich. You are a human. Therefore you are rich.” This inference is logically valid, since in all possible worlds in which all humans were rich, you being a human would be rich, too. That in the actual world not all humans are rich is merely a contingent fact. Of course, in order to apply an inference for purposes of *justification* or *prediction*, one must have good reasons for believing that its premises are true. Valid inferences, however, may also be used for purposes of *falsification*: in such cases, one infers from the falsity of the conclusion that *at least one* of the premises must be false.

To demonstrate that an inference form is semantically *invalid*, it suffices (according to [def. 3.3-2](#)) to find an interpretation that makes all the premises true but the conclusion false. On the other hand, semantic proofs of the *validity* in predicate logic are complicated and abstract. For this reason, logic also uses another procedure: the method of *deductive proofs*. With this method, one assumes some fundamental L-truths as *axioms* and some L-valid inferences as *rules* and aims to derive all other L-true sentences or valid inferences through the iterative application of axioms and rules in the form of “proofs” (see, e.g., [Klenk 1989](#), [Copi et al. 2010](#)). A full logical system consists of a language, a logical semantics, and a deductive proof system (or calculus).

3.4 Conventions of Meaning and Definitional Truths

As with L-true sentences, definitions such as “A bachelor is a hitherto unmarried man” are *analytical* truths, that is, their truth follows from the meaning of their concepts alone, independently of the factual constitution of the world. Differently from L-true sentences, however, the truth of definitions is determined not by the meaning of their logical concepts, but by conventions of meaning for their specific *nonlogical* concepts, in our example “bachelor,” “man,” and “hitherto unmarried.” The truth of sentences which are definitionally true is therefore *not* a matter of their logical form, but it is a matter of the conventions of meaning in the given language. So we define:

(Def. 3.4-1) (a) A sentence is *definitionally true* (or: *extralogically analytically true*) iff its truth is based on certain *conventions of meaning* for its *nonlogical* concepts.

In other words: a sentence is definitionally true iff it follows logically from the set of these conventions of meaning, without being itself logically true.

(b) A sentence is *analytically true* iff it is logically or definitionally true.

(3.4-1) *Examples of definitionally true sentences*:

a. All bachelors are male. $\forall x(Bx \rightarrow Mx)$

b. The length of the standard measure in Paris is one meter. $\text{Length}(a) = (1, m)$

(3.4-2) *Examples of synthetically true sentences:*

- a. All polar bears are white. $\forall x(Px \rightarrow Wx)$
- b. The length of this rod is one meter. $\text{Length}(b) = (1, m)$

The difference between analytic and synthetic truths can be illustrated as follows: if a zoologist sees the truth of (3.4-2)(a) as empirically confirmed, then his justification makes sense. But if someone reported to us that in spite of extensive empirical investigation, he has never met a female bachelor, then this person has clearly not understood the meaning of “bachelor”: bachelors are by definition (unmarried) men; no empirical research is necessary to know this. The examples also make it clear that definitional truths cannot be distinguished from synthetic truths on the basis of logical form, as (3.4-1)(a) and (3.4-2)(a) or (3.4-1)(b) and (3.4-2)(b) have the same logical form, respectively.

Although definitions cannot be *empirically* true or false, they can be more or less *empirically adequate* (see also [Hempel 1965](#), [sec. 6.5](#)). Before Linné, for example, animals and plants were classified by their phenomenological characteristics. Classification by common evolutionary descent, founded by Darwin a century later, turned out to be substantially more adequate, because it reflected a much larger number of common nomological characteristics. Nevertheless, nobody can be forbidden to define “fish” as “water-dwelling vertebrates.” The fact that dolphins, etc., would then also count as fish, although they descended from land-dwelling mammals, with whom they share many more common features than with the other (gill-breathing) fishes, does not disprove this definition – it does however, highlight its empirical inadequacy.

Some authors call particularly natural or empirically adequate definitions ‘true’ or ‘real definitions,’ as distinct from “nominal definitions” in the sense of linguistic conventions. This view goes back to Aristotle and presupposes an *essentialism*, according to which the world holds in itself a ‘true’ partition into natural kinds, which are characterized by certain necessary or “essential” properties (cf. [Fine 1994](#); [Schrenk 2017](#), [sec. 6.3](#)). What one counts as “essential” is however often unclear: there is ongoing debate about the optimal classification of biological taxa. Purely evolutionary classifications are generally held as superior to Linnaean classifications, but they would have to classify crocodiles not together with reptiles, but with birds, which is regarded as ‘unnatural’ by most biological taxonomists (cf. [Ridley 1993](#), 369). This shows that optimal definitions are not only dictated by the facts of the world, but also depend on epistemic matters. For this reason, we do not adopt the notion of ‘real definitions,’ because the demarcation between definitions and synthetic sentences would thereby become *unclear*. By “definitions,” we always understand more or less adequate semantic conventions.

There are three different kinds of sentences which are definitionally true, namely:

(D1) *Explicit definitions*: In these, a concept (the definiendum) is equated by convention with a complex of other concepts (the definiens). For general concepts, an explicit definition has the form of an *equivalence* (see [Suppes 1999](#), sec. 8.3):

(3.4-3)

$\underbrace{\hspace{10em}}_{\text{(Definiendum)}}$		$\underbrace{\hspace{10em}}_{\text{(Definiens)}}$	
$\forall x: \quad x \text{ is a bachelor}$	$\leftrightarrow_{\text{def}}$	$x \text{ is a hitherto unmarried man}$	
<i>Example:</i> $\forall x: \quad x \text{ is a bachelor} \leftrightarrow_{\text{def}} x \text{ is a hitherto unmarried man}$ <i>Logical form:</i> $\forall x(Dx_1 \dots x_n \leftrightarrow A)$, where in A only the individual variables $x_1 \dots x_n$ are free.			

For singular concepts and function concepts, explicit definitions have the form of an identity, for example, $a =_{\text{def}} f(b_1, \dots, b_n)$, or $\forall x(f(x) =_{\text{def}} a \cdot x^n + b)$.

(D2) *Meaning postulates* (also called ‘*implicit*’ definitions) : These are semantic conventions, which like definitions play the role of extralogical-analytic *axioms*, that is, are not arrived at by logical derivation. These meaning postulates, however, do not have the form of an explicit definition, but some other form, mostly that of an implication. *Example:*

$\forall x(x \text{ is round} \rightarrow x \text{ has no corners})$.

(D3) *Derived definitorially true sentences*: These sentences are not logically true, but follow logically from sentences of type D1 or D2.

Example:

$\forall x(x \text{ is not male} \rightarrow x \text{ is not a bachelor})$.

Proper definitions must fulfill the criterion of *eliminability* of defined concepts (cf. Suppes 1957, 154). In order to satisfy this criterion, they have to satisfy the *noncircularity* requirement. For a single explicit definition, this criterion demands that the definiendum concept must not occur in the definiens. If this is the case, then the definiendum of explicit definitions of the form (3.4-3) can be eliminated in every sentence S in which it occurs, simply by replacing it by its definiens: the result is a sentence S* which is definitionally equivalent with S, that is, the equivalence $S \leftrightarrow S^*$ follows from the given definition.² In distinction to noncircular explicit definitions, meaning postulates do *not* fully characterize a concept and do not allow for its elimination.

For an entire *set* of definitions $\{D_1, \dots, D_n\}$, it must be required that this set can be arranged in the form of a *noncircular chain* of definitions, which means that every definiens in this chain contains only concepts which are either undefined or are the definiendum of an *earlier* definition in the chain (cf. von Savigny 1976, 116). An example of a noncircular chain of definitions is $Ax \leftrightarrow_{\text{def}} Bx \wedge \neg Cx$, $Bx \leftrightarrow_{\text{def}} Dx \wedge Ex$, $Dx \leftrightarrow_{\text{def}} \neg Fx$. If we carry out successive replacement in the definiens of Ax , we arrive at $Ax \leftrightarrow_{\text{def}} \neg Fx \wedge Ex \wedge \neg Cx$ as the basic definition of Ax , which contains only primitive (undefined) concepts in its definiens. An example of a circular chain of definitions is $\text{Necessary}(A) \leftrightarrow_{\text{def}} \neg \text{Possible}(\neg A)$, $\text{Possible}(A) \leftrightarrow_{\text{def}} \neg \text{Necessary}(\neg A)$, as after replacement, we arrive at the circular basic definition $\text{Necessary}(A) \leftrightarrow_{\text{def}} \neg \neg \text{Necessary}(\neg A)$.

A further category of definition is *recursive* definitions. They play an important role in logic and computer science. Without going into details, we only mention that in spite of their iterative structure, recursive definitions also satisfy the noncircularity requirement, since they are provably equivalent with (quite complicated) explicit definitions (cf. van Dalen et al. 1978, 157).

The central requirement for analytically true sentences is their empirical non-creativity, which is explicated as follows:

(Prop. 3.4-1) *Requirement of empirical non-creativity*: Analytically true sentences must neither have empirical content, nor create new empirical content in combination with the given system S of accepted background beliefs.

More precisely: A set of sentences Δ can *only* be analytically true if (a) the empirical content of Δ (in the sense of def. 3.7-1) is empty, and (b) the empirical content of S-united-with- Δ is the same as that of S.

Both requirements (a) and (b) must be fulfilled, in order to be able to say that the truth of such sentences is based on analytical convention. In Suppes' criterion (1999, 154) requirement (a) is missing, because he deals only with non-creativity *relative* to S.

Prop. 3.4-1 implies that *no concept may be defined twice* in different ways. It is well known that in the history of measurement, different conventions were suggested. Here are two examples:

1. 1 meter = the length of the Platinum-Iridium-bar in Paris.
2. 1 meter = the length of a pendulum at sea level with an oscillation frequency of one second.

(1) + (2) however, implies the following *empirical consequence*: the length of the Platinum-Iridium-bar in Paris equals the length of a pendulum at sea level with an oscillation frequency of one second. We cannot regard (1) and (2) *simultaneously* as the definition of one meter. If (1) is the accepted definition, then (2) is an empirical hypothesis, and vice versa. Many "definitions" to be found in textbooks, for example, "Water = H_2O ," or "Child = person less than 12 years old," do not fulfill the requirement of empirical non-creativity. This is because the concepts being defined ("water," "child") have already been analytically introduced to common-sense background knowledge in other ways, for example, by ostensive characterizations. Such "definitions" should more rightly be called *explications* in the sense of Carnap (recall sec. 2.1) rather than

definitions. Also so-called “definitions” of theoretical concepts, for example, “force = mass times acceleration,” generate new empirical content and for this reason are not proper definitions (see [sec. 5.1](#)).

As long as a concept in a system of knowledge is characterized only by *one* explicit definition, this definition is provably always empirically non-creative (see [Shoenfield 1967](#), 57f). Also a *single* implicational meaning postulate for a new concept Fx such as $\forall x(Fx \rightarrow Ax)$ is not creative, because it is logically weaker than an explicit definition $\forall x(Fx \leftrightarrow Ax)$. Mostly though, an entire *system* of meaning postulates is introduced for one (or several) concepts; and for this, there is no guarantee of non-creativity. The designation of systems of meaning postulates as “implicit definitions” (which goes back to Hilbert) is therefore easily misunderstood, because “implicit definitions” behave quite differently from proper definitions.

3.5 Classification of Sentences According to Their Generality

The overview is shown in [fig. 3.5-1](#). The logical *strength* decreases successively with the transition from universal via singular to existential sentences: a universal sentence “Everything is an F” is logically stronger than its singular instance “this is an F” which is in turn logically stronger than the existential sentence “something is an F.”

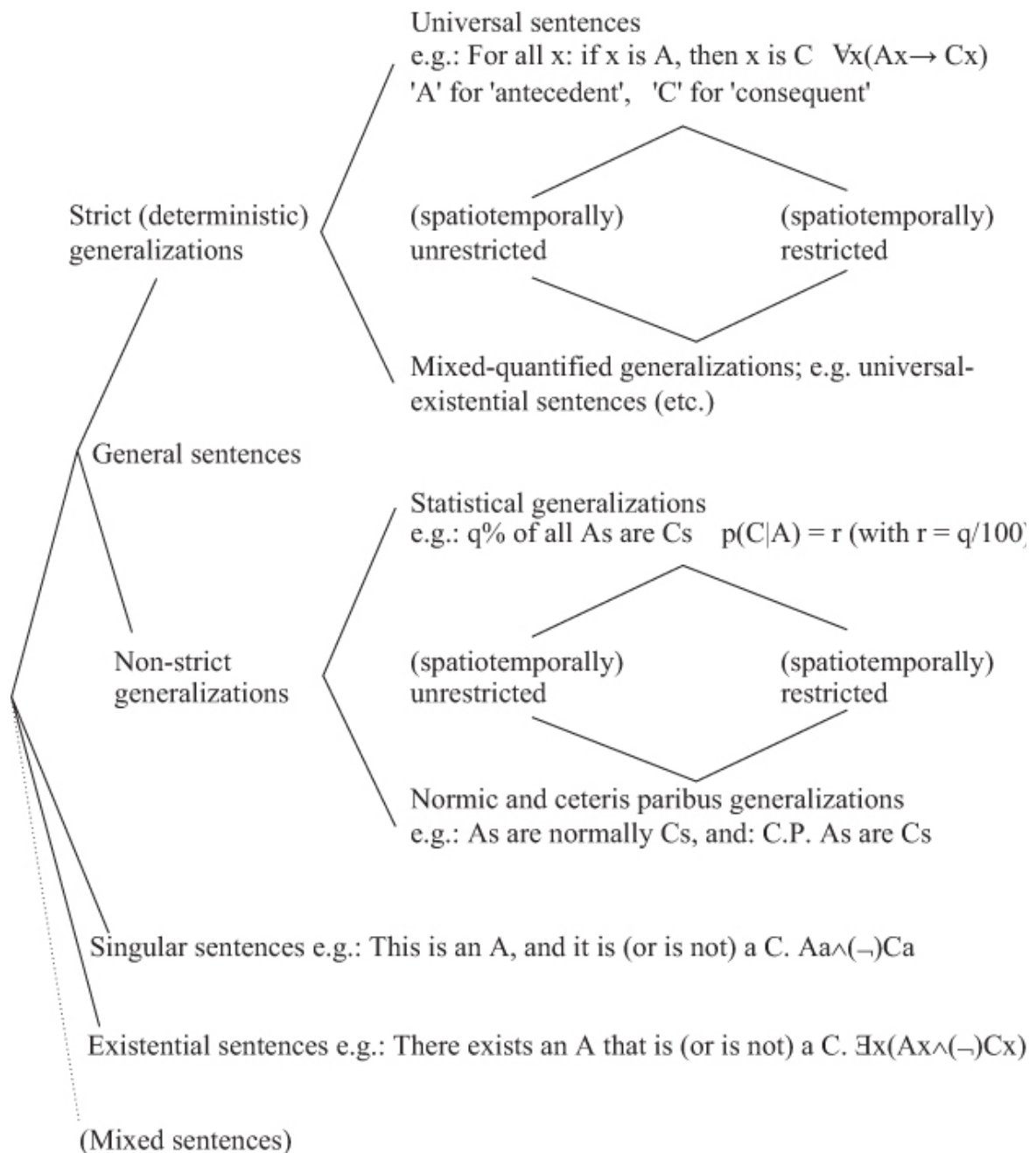


Figure 3.5-1 Classification of sentences according to their generality. ↩

1. *Strict generalizations* contain at least one universal quantifier. They include:

The *universal implication* “For all x: if Ax, then Cx” is the simplest form of a universal sentence. The if-component Ax is also called the *antecedent*, the then-component is called the *consequent*. In general, a *purely* universal sentence begins with one or more universal quantifiers, followed by a quantifier-free formula, which is called the *matrix* of the universal sentence and typically has the form of an implication, an equivalence, or a numerical identity. Universal sentences are also called *strict* or *deterministic* or *exceptionless* generalizations, since the truth of $\forall x(Ax \rightarrow Cx)$ implies for *every* individual a that if Aa obtains, the occurrence of Ca is inevitably determined.

Mixed-quantified generalizations contain both universal and existential quantifiers. *Example*: every action has a motive, formally $\forall x \exists y(\text{Action}(x) \rightarrow \text{Motive}(y, x))$. The meaning of generalizations with mixed

quantifiers depends on the *order* of quantifiers: $\exists x \forall y Rxy$ is a logically stronger statement than $\forall y \exists x Rxy$. *Example:* let “ Rxy ” stand for “ x is the cause of y .” Then $\forall y \exists x Rxy$ says that everything has some cause, but $\exists x \forall y Rxy$ says that some x is the cause of everything, that is, everything has the *same* cause, which is a much stronger statement. If the quantifiers are the same, though, their order is unimportant.

2. *Non-strict generalizations:* In this category, we include primarily *statistical* generalizations of the form “ $q\%$ of all A s are C s.” One calls this a statement of *conditional* probability and writes it down as $p(Cx|Ax) = r$, with $q = r/100$. The probability “ p ” is understood statistically, as a frequency or frequency limit ([sec. 3.9](#)). Although statistical generalizations are general sentences, they are *not* universal sentences. The statistical generalization “ $q\%$ of all A s are C s” says nothing about single individuals of type A , rather it only says something about the class of A s, that is, that $q\%$ of the individuals of the class of A s are located in the class of C s (see [fig. 3.5-2](#)). For this reason, there are no logical consequence relations, but only inductive-probabilistic relations between statistical generalizations and singular sentences, if the domain of individuals is unknown or potentially infinite.

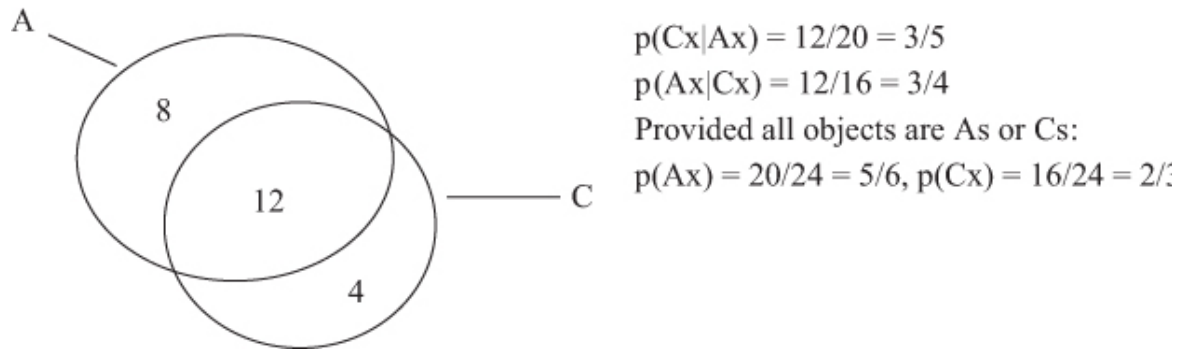


Figure 3.5-2 Statistical generalizations. [↩](#)

A further kind of non-strict generalizations are *normic* generalizations or normal case hypotheses of the form “ A s are *normally* C s” – for example, birds can normally fly (the term “normic” goes back to [Scriven 1959a](#)). For reasons explained in [sec. 4.5.2](#) , we can assume that a normic law analytically *implies* a numerically unspecific statistical generalization of the form “*most* A s are B s,” without being semantically equivalent with it. A third kind of non-strict generalizations are *ceteris paribus* (CP) generalizations, which say that Ax leads to Bx as long as other interfering factors are either *held constant* (comparative CP generalizations), or are *absent* (exclusive CP generalizations, see [sec. 4.5.1](#)).

What statistical, normic, and CP generalizations have in common, is that they allow for *exceptions* – these are instances which verify the antecedent and falsify the consequent predicate. Therefore, inferences from them are described by a non-monotonic consequence relation in the sense of [sec. 2.7.4](#) . Strict laws forbid such exceptions and are falsified by them.

Both strict and non-strict generalizations can be unrestricted or restricted in space and time. They are restricted if the antecedent predicate contains a restriction to a finite area of space and time – for example, “all swans in *Europe* are white,” as opposed to “all swans are white.” Unrestricted universal sentences are also called *spatiotemporally universal* ([sec. 6.5.4](#)).

3. *Singular* are all sentences which contain *no quantifiers* , and also no statistical probability function. Singular sentences are composed of atomic sentences using propositional operators (Fa , $\neg Fa$, $Fa \wedge Rab$, etc.). They always say something about one or more individual things.

4. *Existential sentences:* Pure existential sentences begin with one or more existential quantifiers, followed by a quantifier-free matrix formula. Usually, the matrix has the form of a conjunction $\exists x(Ax \wedge \pm Cx)$; for example, there is a black swan. Such existential sentences may also be unrestricted or restricted, depending on whether Ax contains a restriction in space and time or not.

5. *Mixed sentences* , finally, are propositional combinations of the above sentence types, like for instance $\forall x(Fx \rightarrow Gx) \vee Ha$. They play a subordinate role.

The table in [prop. 3.5-1](#) summarizes the most important logico-deductive relationships between the listed kinds of sentences.

(Prop. 3.5-1) *Notation*: “ $X \Vdash Y$ ” stands for “ Y follows logically from X ” (“ $\dashv$ ” for “logically equivalent”).

1. Universal sentence $\Vdash$ singular sentence
All As are Cs $\Vdash$ if a is A, then a is C
2. Universal sentence and singular sentence $\Vdash$ singular sentence (*explanation scheme*)
All As are Cs, and a is A $\Vdash$ a is C
3. Singular sentence falsifies strictly universal sentence (*falsification scheme I*)
Singular sentence $\Vdash$ negation of a universal sentence
a is A and not C $\Vdash$ not all As are Cs
4. Existential sentence falsifies strictly universal sentence (*falsification scheme II*)
There is an A that is not a C $\Vdash$ not all As are Cs
5. Universal sentence is logically equivalent to the negation of an existential sentence.
All As are Cs $\dashv$ $\dashv$ there is no A that is not a C
6. Singular sentence $\Vdash$ Existential sentence
a is an A $\Vdash$ there is an x that is an A
7. There are no logical consequence relations between non-strict generalizations and singular sentences, only statistical or epistemic relations of probability.

The above characterizations of kinds of sentences are of a *syntactic* nature. Syntactic characterizations are not always invariant with respect to (w.r.t.) logically equivalent (L-equivalent) transformations. For instance, the singular sentence Fa is L-equivalent to the universal sentence $\forall x(x = a \rightarrow Fx)$ (all x which are identical with a have property F). This sentence, however, is not a universal sentence in the *semantic* sense, as it only says something about the individual a . If characterizations of sentences should express genuine properties of the *propositions* thereby expressed, they must be invariant w.r.t. L-equivalent sentence transformation, since L-equivalent sentences express the same proposition. A standard method of solving this problem is the method of *essential* sentence types (see [Hempel/Oppenheim 1948](#) , 271). ³ Here, one regards a sentence as essentially universal (or existential) only if it is not L-equivalent to a singular sentence:

(Def. 3.5-1) *Essential sentence types (according to their generality)*:

1. A sentence S is an essentially universal/existential sentence iff S is a syntactically universal/existential sentence and not L-equivalent to a syntactically singular sentence.
2. By contrast, a sentence S is an *essentially* singular sentence iff S is L-equivalent to a syntactically singular sentence.

The division of kinds of sentences according to their generality is largely *independent* of the content-based classification of sentences given in [sec. 3.2](#) : the kinds of sentences listed above can be either empirical or theoretical sentences, depending on which nonlogical concepts occur in these sentences, and which range their quantifiers possess. Unrestricted universal or existential sentences can however never be observation sentences.

A further distinction, which runs *crosswise* to our distinctions up to now, is the one between *evidences* and *hypotheses* . It relates to the *epistemic status* of our knowledge. Evidence statements are very well-secured sentences – actual observation sentences are prototypical for this category. Hypotheses are less well-secured sentences, which are in need of further testing – prototypical for this are general sentences, or potential observation sentences (i.e. observable predictions).

PHILOSOPHY OF SCIENCE

A Unified Approach

GERHARD SCHURZ

Second Edition



PHILOSOPHY OF SCIENCE

A Unified Approach

GERHARD SCHURZ

Second Edition



“I consider this book to be a unique work both with regard to the sheer amount of material it covers but also with regard to the clarity, depth and insight with which these matters are dealt with by one of Europe’s foremost analytic philosophers. I would not hesitate to use it in an advanced course on the subject, and I believe it would work well also for an introductory course given the careful way in which the book is structured into introductory and advanced material.”

– **Erik J. Olsson** , *Lund University*

“ *Philosophy of Science: A Unified Approach* revamps and bolsters the logical empiricist way to think about and to practice philosophy of science. It combines technical precision and rigor with philosophical breadth and depth. It delves into all major philosophical problems and issues in general philosophy of science. It elegantly presents the received positions and approaches, but it also articulates novel views and arguments. It systematically explains the necessary technical concepts and makes them do deep philosophical work. *Philosophy of Science: A Unified Approach* lives up to its promise to highlight and defend the methodological unity of the sciences. Gerhard Schurz has written a superb introduction to philosophy of science; novices and experts alike have a lot to learn from it.”

Stathis Psillos , *University of Western Ontario and University of Athens*

“This is an excellent introduction to the general philosophy of science. It is lucidly written and very well structured in elementary and advanced subsections. Inspired by the classic authors of the 20th century, such as Carnap and Popper, and guided by the methodological version of the idea of the unity of the sciences and the humanities, Gerhard Schurz shows how it stands with the main topics in the first decades of the 21st century.”

– **Theo Kuipers** , *University of Groningen*

Philosophy of Science

Philosophy of Science: A Unified Approach, Second Edition, combines a general introduction to philosophy of science with an integrated survey of all its important subfields. As this book's subtitle suggests, this excellent overview is guided methodologically by "a unified approach" to philosophy of science. Behind the diversity of scientific fields, one can recognize a methodological unity. This unity is worked out in this book, revealing all the while important differences between subject areas.

This book offers a two-part approach that makes it an excellent introduction for readers new to the field and a useful resource for more advanced students. Each chapter is divided into two sections. The first section assumes no foreknowledge of the subject introduced, and the second section builds upon the first by bringing into the conversation more advanced, complementary topics from contemporary research.

Definitions, key propositions, examples, and figures complement all of the core material. At the end of every chapter, there are selected readings and exercise questions (with solutions and answers to many of them at the end of this book). This book also includes a comprehensive bibliography and an index.

The Second Edition has been revised to make the technical material much more accessible for students outside of the sciences and philosophy. It also incorporates new research over the last decade, including significantly expanded coverage of values in science, mechanistic explanations, and inductive risk. An extended index of subjects that can also be used as a glossary has been added to the back of this book.

Gerhard Schurz (MSc 1980, PhD 1983) is Professor of Philosophy at the Heinrich Heine University Düsseldorf (since March 2002) and member of the Leopoldina (German Academy of Sciences) and the International Academy for Philosophy of Science. Before 2002, he was an Associated Professor at the University of Salzburg and a Guest Professor at the University of California at Irvine and Yale University. His research areas include philosophy of science, epistemology, logic, and cognitive science. He is the author of 270 research articles and 12 books.

Philosophy of Science

A Unified Approach

Second Edition

Gerhard Schurz



Designed cover image: Getty Images

Second edition published 2026
by Routledge
605 Third Avenue, New York, NY 10158

and by Routledge
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2026 Gerhard Schurz

The right of Gerhard Schurz to be identified as author of this work has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

The Open Access version of this book, available at www.taylorfrancis.com, has been made available under a Creative Commons Attribution-ShareAlike 4.0 (CC BY-SA) International license.

Any third party material in this book is not included in the OA Creative Commons license, unless indicated otherwise in a credit line to the material. Please direct any permissions enquiries to the original rightsholder.

For Product Safety Concerns and Information please contact our EU representative
GPSR@taylorandfrancis.com. Taylor & Francis Verlag GmbH, Kaufingerstraße 24, 80331 München, Germany.

Trademark notice : Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

First edition published by Routledge 2013

Library of Congress Cataloging-in-Publication Data

Names: Schurz, Gerhard, 1956– author

Title: Philosophy of science : a unified approach / Gerhard Schurz.

Other titles: Einführung in die Wissenschaftstheorie. English

Description: 2nd edition. | New York, NY ; Abingdon, Oxon : Routledge, 2026. | Includes bibliographical references and index. | Text in English. Translation from German.

Identifiers: LCCN 2026005541 | ISBN 9781041145820 hardback | ISBN 9781041145813 paperback | ISBN 9781003675075 ebook

Subjects: LCSH: Science—Philosophy

Classification: LCC Q175 .S413414 2026

LC record available at <https://lcn.loc.gov/2026005541>

ISBN: 978-1-041-14582-0 (hbk)

ISBN: 978-1-041-14581-3 (pbk)

ISBN: 978-1-003-67507-5 (ebk)

DOI: [10.4324/9781003675075](https://doi.org/10.4324/9781003675075)

Typeset in Sabon
by codeMantra

Contents

[List of Figures, Definitions, and Key Propositions](#)

[List of Symbols](#)

[Foreword](#)

[Foreword to the Second Edition](#)

[1 Introduction: Where Do We Stand?](#)

- [1.1 Tasks and Aims of Philosophy of Science](#)
- [1.2 On the Historical Development of Philosophy of Science](#)
- [1.3 Philosophical Positions in Contemporary Philosophy of Science](#)
 - [1.3.1 Logical Empiricism](#)
 - [1.3.2 Critical Rationalism](#)
 - [1.3.3 The Historical Account and Relativism](#)
 - [1.3.4 Metaphysical and Scientific Realism](#)
 - [1.3.5 Naturalism and the Cognitive Turn](#)
 - [1.3.6 Structuralist and Semantic View of Theories](#)
 - [1.3.7 Pragmatic Philosophy of Science](#)
 - [1.3.8 Hermeneutics and Critical Theory](#)
 - [1.3.9 Radical Constructivism, Postmodernism, and the Science Wars](#)
- [1.4 Introductory Literature and Exercises](#)

[2 The Question of Unity: In Search for Common Foundations of the Sciences](#)

- [2.1 Normative or Descriptive? The Method of Rational Reconstruction](#)
- [2.2 Common Epistemological Assumptions](#)
- [2.3 Common Methodological Features](#)
- [2.4 Scientific Disciplines and Their Classification](#)
 - [2.4.1 The Special Status of Mathematical Sciences](#)
 - [2.4.2 Testing the Common Foundations of Science](#)
 - [2.4.3 Two Further Classifications of Sciences](#)
- [2.5 The Question of Value Neutrality](#)
 - [2.5.1 Max Weber's Value-Freedom Postulate](#)
 - [2.5.2 Value Neutrality and Means-End Inferences](#)
 - [2.5.3 Explication of the Value-Neutrality Requirement](#)
- [2.6 The Demarcation Problem and the Methodological Unity of Sciences](#)
 - [2.6.1 Demarcation](#)

2.6.2 Methodological Unity instead of Physical Reductionism

2.7 Scientific Inference and Argumentation

2.7.1 Deduction and Induction

2.7.2 Popper and the Significance of Inductive Inference in the Sciences

2.7.3 Abduction and Inference to the Best Explanation

2.7.4 Monotonic and Non-monotonic Inferences

Complementary and Advanced Topics

2.8 How Much Metaphysics is Presupposed by Science?

2.8.1 Truth, Idealism, and Realism

2.8.2 Constructive Realism versus Radical Constructivism

2.9 On the Theory (In)dependence of Observation

2.9.1 Theory-Dependency of Observation – Pros and Cons

2.9.2 A Definition of Theory-Neutral Observation Concepts

2.10 Value Neutrality in the Context of Contemporary Debates

2.10.1 The Role of Values in Science

2.10.2 The Argument of Inductive Risk

2.10.3 Meta-ethical Justification of Value Neutrality and the Difference between Experience and Valuation

2.11 The Justification of Induction: An Unsolvable Problem?

2.11.1 Is Induction Rationally Unjustifiable?

2.11.2 The Optimality of Meta-induction

2.12 Introductory Literature and Exercises

3 The Conceptual Toolkit: Language, Logic, and Probability

3.1 Kinds and Classification of Concepts

3.1.1 The Logical Type of Concepts

3.1.2 Syntax and Semantics, Statements, and Models

3.1.3 The Content Type of Concepts

3.1.4 The Gradation (Scale) Type of Concepts

3.2 Classification of Sentences According to Their Content Type

3.2.1 Descriptive versus Prescriptive Sentences

3.2.2 Observation and Basic Sentences, Empirical and Theoretical Sentences

3.3 Logical Truth and Deductive Logic

3.4 Conventions of Meaning and Definitional Truths

3.5 Classification of Sentences According to Their Generality

3.6 General Sentences, Lawlikeness, Determinism, and Indeterminism

3.7 Content of Sentences and Kinds of Content

3.8 Verification, Falsification, Confirmation, and Disconfirmation

3.9 Objective (Statistical) and Subjective (Epistemic) Probability

Complementary and Advanced Topics

3.10 Disposition Concepts

3.11 Challenges to the Logical-Definitional-Synthetic Distinction

- [3.12 Relevance and Irrelevance in Logical Inferences](#)
- [3.13 More on Probability](#)
 - [3.13.1 Mathematical Laws of Probability](#)
 - [3.13.2 Problems of Objective-Statistical Probability Concepts](#)
 - [3.13.3 Problems of Subjective Probability Concepts](#)
 - [3.13.4 Connections between Objective and Subjective Probabilities: Principal Principle and Narrowest Reference Class](#)
- [3.14 Introductory Literature and Exercises](#)

4 A Question of Fit: Law Hypotheses and Their Empirical Testing

- [4.1 The Condition of Relevance](#)
 - [4.1.1 Relevance of Strict Laws](#)
 - [4.1.2 Relevance of Statistical Laws](#)
- [4.2 Testing of Strict Law Hypotheses for Truth and Relevance](#)
 - [4.2.1 The Method of Agreement and Difference](#)
 - [4.2.2 Methodological Induction: How to Discover Strict Law Hypotheses](#)
- [4.3 Testing Statistical Laws](#)
 - [4.3.1 Testing for Likely Truth and Acceptance Intervals](#)
 - [4.3.2 Discovery of Statistical Hypotheses and Confidence Intervals](#)
 - [4.3.3 Test Statistics and Inference Statistics](#)
 - [4.3.4 Testing for Relevance and Significant Differences](#)
 - [4.3.5 Statistical Representativeness](#)
 - [4.3.6 Sources of Error in Statistical Methods](#)
 - [4.3.7 Applying Statistical Hypotheses to Individual Cases](#)
- [4.4 Correlation and Causality](#)
 - [4.4.1 Hidden Variables](#)
 - [4.4.2 Causal Direction](#)

Complementary and Advanced Topics

- [4.5 Ceteris Paribus Laws and Normic Laws](#)
 - [4.5.1 Comparative versus Exclusive CP Laws](#)
 - [4.5.2 Normic Laws and Their Evolution–Theoretic Foundation](#)
- [4.6 Probability Distributions for Continuous Variables](#)
- [4.7 From Classical Statistics to Bayesian Statistics](#)
 - [4.7.1 The Likelihood Intuition and Its Bayesianism Justification](#)
 - [4.7.2 Objective Bayesianism and the Principle of Indifference](#)
 - [4.7.3 Subjective Bayesianism and Convergence of Opinions](#)
- [4.8 Introductory Literature and Exercises](#)

5 Going beyond Experience: Theories and Their Empirical Evaluation

- [5.1 Theoretical Concepts and Multiple Laws of Correspondence](#)
- [5.2 The Example of Newtonian Physics](#)
- [5.3 The Structure of Scientific Theories](#)
 - [5.3.1 Components of Theories](#)
 - [5.3.2 Theory Nets, Intertheoretical Relations, and Kinds of Theoretical Concepts](#)

- 5.4 [Methodological Features of \(Good\) Scientific Theories](#)
 - 5.4.1 [System Character: Holism of Meaning, Content, and Theory Testing \(Duhem's Thesis\)](#)
 - 5.4.2 [Empirical Creativity, Global Unification, and Novel Predictions: Answers to Hempel's Dilemma](#)
- 5.5 [The Example of Piaget's Theory of Cognitive Development](#)
- 5.6 [The Dynamics of Scientific Theories](#)
 - 5.6.1 [Lakatos' Model of Theory Revision](#)
 - 5.6.2 [Theory Evaluation and Theory Progress](#)
- 5.7 [Instrumentalism and Realism: The Ontology of Scientific Theories](#)

[Complementary and Advanced Topics](#)

- 5.8 [Instrumentalism-Realism Controversies in Philosophy of Science](#)
 - 5.8.1 [The Ramsey Sentence of a Theory: An Instrumentalist Challenge](#)
 - 5.8.2 [The Realist Interpretation of the Ramsey Sentence](#)
 - 5.8.3 [The No-Miracles Argument](#)
 - 5.8.4 [The Pessimistic Meta-Induction](#)
 - 5.8.5 [The Empirical Underdetermination of Theories](#)
 - 5.8.6 [Intertheoretical Correspondence, Structural and Selective Realism](#)
 - 5.8.7 [Criteria for Empirical Significance](#)
 - 5.8.8 [Probabilistic Analysis of the NMA, Alternative Theories, and Local Realism](#)
- 5.9 [General Accounts of Confirmation](#)
 - 5.9.1 [Hypothetico-Deductive Confirmation](#)
 - 5.9.2 [Probabilistic Confirmation and Bayesianism](#)
 - 5.9.3 [Bayesian Pseudo-Confirmation through Content Cutting versus Genuine Confirmation](#)
 - 5.9.4 [Goodman's Paradox](#)
 - 5.9.5 [Confirmation and Acceptance](#)
- 5.10 [Testing Theories: Parameter Fitting and Use-Novel Evidence](#)
 - 5.10.1 [Post-Factum Speculations](#)
 - 5.10.2 [Curve Fitting](#)
- 5.11 [Non-Confirmational Accounts of Theory Evaluation](#)
 - 5.11.1 [Truthlikeness](#)
 - 5.11.2 [Unification, Coherence, and Simplicity](#)
- 5.12 [The Problem of Language Dependence](#)
- 5.13 [Introductory Literature and Exercises](#)

[6 In Search of Causes: Explanation and All That Goes with It](#)

- 6.1 [The Deductive-Nomological Model of Explanation](#)
- 6.2 [Explanation versus Prediction and Justification](#)
 - 6.2.1 [Predictions without Explanatory Function](#)
 - 6.2.2 [Causality and Lawlikeness](#)
 - 6.2.3 [Explanations without Predictive Function](#)
 - 6.2.4 [Logical Problems of Irrelevance and Redundancy](#)

6.3 Probabilistic Explanation Models

6.3.1 Inductive-Statistical Explanation

6.3.2 Maximal Specificity

6.3.3 Probabilistic and Causal Relevance Models

6.3.4 Conflicting Intuitions about the Magnitude of Explanatory Probabilities

6.4 Further Types of Explanation

6.4.1 Mechanistic Explanations

6.4.2 Normic Explanations and the Explanation of Human Actions

6.4.3 Explanation and Understanding as Unification

Complementary and Advanced Topics

6.5 Lawlikeness

6.5.1 Laws of Nature versus System Laws

6.5.2 Lawlikeness in the Wide versus Narrow Sense

6.5.3 Lawlikeness i.w.s. and Counterfactuals

6.5.4 Lawlikeness i.n.s., Spatiotemporal Universality, and Maxwell's Condition

6.5.5 The Best System Account

6.5.6 Physical Necessity and Independent Possibility Knowledge

6.6 Causality

6.6.1 Singular and General Causality

6.6.2 Regularity Accounts

6.6.3 Counterfactual Accounts

6.6.4 Causation and the Direction of Time

6.6.5 Causal Process Theories

6.6.6 Interventionist Accounts

6.6.7 Causality as a Theoretical Concept

6.6.8 The Theory of Causal Nets

6.6.9 Causal Discovery and Empirical Content of TCN

6.6.10 From Variables to Events: Overdetermination and Causal Preemption

6.7 Introductory Literature and Exercises

Bibliography

Solutions to Selected Exercises

Index of Names

Index of Subjects

List of Figures, Definitions, and Key Propositions

Figures

- [2.1-1 The method of rational reconstruction](#)
- [2.3-1 The three levels of the scientific method](#)
- [2.4-1 Classification of the factual sciences](#)
- [2.5-1 Schematic representation of the value neutrality requirement](#)
- [2.7-1 The inductive-deductive scheme](#)
- [2.7-2 Interaction of epistemic induction and abduction](#)
- [2.9-1 Duck-rabbit](#)
- [2.9-2 Mueller-Lyer illusion](#)
- [2.9-3 Expected learning curves in ostensive learning experiments](#)
- [3.1-1 Division of concepts according to their logical type](#)
- [3.1-2 Classification of concepts according to their content type](#)
- [3.1-3 Classification of concepts according to their gradation type](#)
- [3.1-4 Ordinal scale](#)
- [3.2-1 Classification of sentences according to their content type](#)
- [3.5-1 Classification of sentences according to their generality](#)
- [3.5-2 Statistical generalizations](#)
- [3.8-1 Testability of generalizations](#)
- [3.9-1 Non-monotonicity of conditional probabilities](#)
- [3.13-1 Three binomial distributions](#)
- [4.1-1 Irrelevance and joke](#)
- [4.3-1 Acceptance interval](#)
- [4.3-2 Acceptance and confidence interval](#)
- [4.3-3 Probability distribution of sample differences and significant difference](#)
- [4.4-1 Common cause](#)
- [4.4-2 Common cause vs. intermediate cause](#)
- [4.4-3 Spurious causal independence](#)
- [4.4-4 What is the direction of causality?](#)
- [4.6-1 Normal distribution](#)
- [4.7-1 Shortest interval](#)
- [4.7-2 Transformation of uniform into non-uniform density](#)
- [5.1-1 One theoretical feature causes many empirical dispositions](#)
- [5.1-2 Beam balances and spring scales on Earth and Uranus](#)
- [5.2-1 Theory lattice in classical mechanics](#)

[5.3-1,2,3 Subdivisions of a theory](#)
[5.3-4 Theory net of physical and chemical theories](#)
[5.6-1 Historical theory sequence](#)
[5.6-2 Theory comparison](#)
[5.6-3 Empirically complementary theories](#)
[5.6-4 Theories of optics after Bunge](#)
[5.10-1 Curve fitting and use-novel evidence](#)
[6.6-1 Explanation of screening-off](#)
[6.6-2 Three statistically strongly indistinguishable graphs](#)
[6.6-3 Causal graph of variables and related graphs of values](#)

Definitions

[\(Def. 2.1-1\) Method of rational reconstruction](#)
[\(Def. 2.5-1\) The general scheme of means-end inferences](#)
[\(Def. 2.7-1\) Monotonicity](#)
[\(Def. 2.9-1\) Theory-neutral observation concepts](#)
[\(Def. 3.2-1\) Descriptive versus prescriptive sentences](#)
[\(Def. 3.2-2\) Observational, empirical, theoretical and T-theoretical sentences](#)
[\(Def. 3.3-1\) Logical truth](#)
[\(Def. 3.3-2\) Logical validity](#)
[\(Def. 3.4-1\) Definitional truth](#)
[\(Def. 3.5-1\) Essential sentence types](#)
[\(Def. 3.7-1\) Logical and empirical content](#)
[\(Def. 3.8-1\) Verification, falsification, confirmation and disconfirmation](#)
[\(Def. 3.9-1\) Statistical and subjective probability](#)
[\(Def. 3.9-2\) Conditional probability](#)
[\(Def. 3.9-3\) Principle of narrowest reference class](#)
[\(Def. 3.10-1\) Bilateral reduction sentence](#)
[\(Def. 3.10-2\) Nomological definition of dispositions](#)
[\(Def. 3.12-1\) Relevance in logical inferences](#)
[\(Def. 3.12-2\) Content elements](#)
[\(Def. 3.12-3\) Relevant content](#)
[\(Def. 3.12-4\) Irreducible representation](#)
[\(Def. 4.1-1\) Condition of relevance for strict law hypotheses](#)
[\(Def. 4.1-2\) Statistical relevance for conjunctive antecedents](#)
[\(Def. 4.3-1\) Acceptance interval](#)
[\(Def. 4.3-2\) Significant difference](#)
[\(Def. 4.6-1\) Mean, variance and standard deviation](#)
[\(Def. 5.3-1\) Kinds of theoretical concepts](#)
[\(Def. 5.8-1\) Ramsey sentence](#)
[\(Def. 5.8-2\) Carnap sentence](#)
[\(Def. 5.8-3\) Ramsey-Lewis sentence](#)
[\(Def. 5.9-1\) Classical HD-confirmation](#)
[\(Def. 5.9-2\) Relevant HD-confirmation](#)
[\(Def. 5.9-3\) Incremental and absolute Bayesian confirmation](#)
[\(Def. 5.9-4\) Genuine confirmation \(full and partial\)](#)
[\(Def. 5.9-5\) Goodman's grue predicate](#)
[\(Def. 5.11-1\) Popper-truthlikeness](#)
[\(Def. 5.11-2\) Schurz/Weingartner-truthlikeness](#)

[\(Def. 6.5-1\) Spatiotemporal universality](#)
[\(Def. 6.6-2\) Counterfactual Causality according to Lewis](#)
[\(Def. 6.6-3\) Interventionist causality after Woodward](#)
[\(Def. 6.6-4\) Screening-off between variables](#)
[\(Def. 6.6-5\) Linking-up between variables](#)
[\(Def. 6.6-6\) Causal Markov axiom](#)
[\(Def. 6.6-7\) Faithfulness condition](#)
[\(Def. 6.6-8\) Minimality and productivity](#)
[\(Def. 6.6-9\) Cause-effect relation between values of variables](#)

Key Propositions

[\(Prop. 2.1-1\) The supreme epistemic goal of science](#)
[\(Prop. 2.5-1\) Value neutrality requirement](#)
[\(Prop. 2.7-1\) General pattern of abduction](#)
[\(Prop. 2.7-2\) Inference to the best explanation](#)
[\(Prop. 2.7-3\) Application rule for monotonic inferences](#)
[\(Prop. 2.7-4\) Application rule for non-monotonic inferences](#)
[\(Prop. 2.11-1\) Universal optimality of AW](#)
[\(Prop. 3.1-1\) Translation between logical and set-theoretical language](#)
[\(Prop. 3.1-2\) Linguistic and model-theoretic theory representation](#)
[\(Prop. 3.4-1\) Requirement of empirical non-creativity](#)
[\(Prop. 3.5-1\) Important logical relationships](#)
[\(Prop. 3.9-1\) Basic axioms of probability](#)
[\(Prop. 3.13-1\) Theorems of unconditional probability](#)
[\(Prop. 3.13-2\) Theorems of conditional probability](#)
[\(Prop. 3.13-3\) Binomial formula](#)
[\(Prop. 3.13-4\) Subjective probabilities as coherent fair betting quotients](#)
[\(Prop. 3.13-5\) Statistical principal principle](#)
[\(Prop. 3.13-6\) Weak induction principles for subjective probabilities](#)
[\(Prop. 4.1-1\) Statistical relevance and correlation for qualitative properties](#)
[\(Prop. 4.2-1\) Testing a strict law hypothesis](#)
[\(Prop. 4.3-1\) Testing a statistical law hypothesis](#)
[\(Prop. 4.4-1\) Reichenbach conditions for screening off](#)
[\(Prop. 4.5-1\) Comparative CP law](#)
[\(Prop. 4.5-2\) Exclusive CP assertion](#)
[\(Prop. 4.5-3\) Normic interpretation of eCP laws](#)
[\(Prop. 4.6-1\) Mean value, variance and standard deviation of the distribution of sample means](#)
[\(Prop. 4.7-1\) Principle of Bayes-statistics](#)
[\(Prop. 4.7-2\) Bayesian justification of the likelihood-intuition](#)
[\(Prop. 4.7-3\) Consequences of the indifference principle](#)
[\(Prop. 4.7-4\) Bayesian convergence results](#)
[\(Prop. 4.7-5\) Biased priors](#)
[\(Prop. 5.1-1\) Synthetic nature of laws of correspondence for theoretical concepts](#)
[\(Prop. 5.1-2,3\) Biased indicators](#)
[\(Prop. 5.4-1\) Logical background of the holism of testing](#)
[\(Prop. 5.4-2\) Empirical creativity](#)
[\(Prop. 5.6-1\) Falsification of theory version](#)
[\(Prop. 5.6-2,3\) Successes and failures of theory versions](#)
[\(Prop. 5.6-4\) Confirmation of theory versions](#)

[\(Prop. 5.6-5\) Dynamical truthlikeness](#)
[\(Prop. 5.6-6\) Theory comparison and theory progress](#)
[\(Prop. 5.6-7\) Rational acceptability of theory versions](#)
[\(Prop. 5.6-8\) Untested empirical excess content](#)
[\(Prop. 5.6-9\) Confirmation of theory cores](#)
[\(Prop. 5.6-10\) Preference between theory cores](#)
[\(Prop. 5.6-11\) Rational acceptability of theory cores](#)
[\(Prop. 5.8-1\) Newman's problem](#)
[\(Prop. 5.8-2\) Howson's analysis of the NMA](#)
[\(Prop. 5.9-1\) HD-confirmation and Bayesian confirmation](#)
[\(Prop. 5.9-2\) Impossibility results for rational acceptance](#)
[\(Prop. 5.10-1\) Necessary criteria for transfer of probability increase](#)
[\(Prop. 6.1-1\) DN-explanation of events](#)
[\(Prop. 6.1-2\) DN-explanation of laws](#)
[\(Prop. 6.2-1\) Lawlikeness and causality in DN-explanations](#)
[\(Prop. 6.2-2\) Predictive function](#)
[\(Prop. 6.2-3\) Logical relevance conditions for DN-explanations](#)
[\(Prop. 6.3-1\) Inductive-statistical explanation](#)
[\(Prop. 6.3-2\) Maximal specificity](#)
[\(Prop. 6.4-1\) Normic explanation](#)
[\(Prop. 6.4-2\) Unification condition](#)
[\(Prop. 6.5-1\) Physical necessity](#)
[\(Prop. 6.6-1\) Singular and general causality](#)
[\(Prop. 6.6-2\) Two versions of the causal Markov axiom](#)
[\(Prop. 6.6-3\) Equivalence of productivity and minimality](#)
[\(Prop. 6.6-4\) Causal discovery algorithm](#)
[\(Prop. 6.6-5\) Empirical content of TCN](#)
[\(Prop. 6.6-6\) Causal nets with time-indexed variables](#)

List of Symbols

$\neg$	not
$\wedge$	and
$\vee$	or
$\rightarrow$	material implication (if-then)
$\leftrightarrow$	equivalent (if-and-only-if)
$\leftrightarrow_{\text{def}}$	equivalent by definition
$\forall x$	for all x
$\exists x$	there exists an x
$\exists!x$	there exists exactly one x
$=$	identical
$\neq$	not identical
$=_{\text{def}}$	identical by definition
$\square$	necessary
$\mathbb{N}$	set of natural numbers
$\mathbb{R}$	set of real numbers
$\in$	is an element of

$\{a_1, \dots, a_n\}$	set of individuals
$(a_1, \dots, a_n)$	ordered sequence of individuals
$\{x: Fx\}$	set of all individuals which possess property F
$\subseteq$	proper or improper subset
$\subset$	proper subset
$\cap$	intersection
$\cup$	union
$-$	applied to sets set-theoretical complement
$\times$	Cartesian product between sets
$\cdot$	ordinary product between numbers
$/$	division (example $x \cdot y / z = x$ times y divided through z)
$\square$	empty set
$\sum_i \in_I x_i$ and $\prod_i \in_I x_i$	sum and product (respectively) over all x_i with i in I
D	domain
D^n	n -fold Cartesian product of D
I	interpretation function
$\mathcal{L}$	formal language
$M(\mathcal{L})$	set of models for language $\mathcal{L}$
$\leq$	smaller or equal
$\geq$	greater or equal
$\equiv$	equivalence relation

$f:A \rightarrow B$	function from domain A into set of values B
$\models$	verifies (makes true)
$\Vdash$	follows logically from
$\nVdash$	does not follow logically from
$\dashv \vdash$	is logically equivalent with
p	statistical probability
P	epistemic probability
$C(S)$ and $C_r(S)$	logical content and relevant logical content of a sentence set S
$E(S)$ and $E_r(S)$	empirical content and relevant empirical content of a sentence set S
Ω	possibility space
AL	algebra
$\ A\ $	semantic extension of formula A
$df(x)/dx$	first derivative of f(x)
$\sqrt{}$	square root
$\int$	integral
$\lim_{n \rightarrow \infty}$	limit for n going to infinity
$R(T)$	Ramsey sentence of theory T
$(IN) \text{ DEP } (X,Y)$	variables X and Y are probabilistically (in)dependent

Foreword

I have written *Philosophy of Science: A Unified Approach* for beginners as well as for graduate students and experts. This twofold orientation is reflected in the major structure of the chapters, which start with introductory sections for beginners, followed by sections entitled “complementary and advanced topics,” in which these topics are explained in a condensed way, including presentations of my own accounts. In this way, this book attempts to offer not only an introduction, but also a detailed overview of the various areas of the Philosophy of Science and their interrelations, which will help the reader familiarize him- or herself with the present state of the art in this field.

This book addresses itself both to philosophers and to scientists of all disciplines. By way of bridging Philosophy of Science with neighboring fields such as statistics, hermeneutics, cognitive science, the debate on science and values, and detailed examples of contemporary theories in sciences and humanities, this book tries to convey a deep understanding of the sciences and their methods. This book is *self-contained* and written without presuppositions about the readers. On rare occasions, formulas are used, but I’ve always explained them in a clear and easy-to-follow language.

Although this book introduces the reader to all major areas of philosophy of science, it also has a bias and a general philosophical message. Concerning the bias, the author understands himself as a follower of logically and empirically oriented philosophers of science such as Carnap, Hempel, Reichenbach, and Stegmüller in the German-speaking area. This book’s general message is reflected in its subtitle, *A Unified Approach*. Behind the diversity of phenomena investigated in disciplines as different as elementary particle physics and social sciences, or chemistry and linguistics, one can recognize a deep common ground that unites all of the sciences. This unity, however, is not so much a unity at the level of objects (as it was claimed by the proponents of physical reductionism), but rather a *methodological unity*. [Chapter 2](#) is concentrated on finding the common methodological features of the sciences and their epistemological presuppositions. The remaining chapters will not only work out these common features in all details but will also reveal some important differences between characteristic groups of disciplines, such as the physical sciences versus special sciences (e.g. biology, psychology, or social sciences).

Use in teaching: By restricting the teaching materials to Chapters 1–5 and omitting all sections on advanced topics, one obtains the basis for a semester-long two-hours-per-week course. Covering all of the material contained in this book will fill two such courses. The *core materials* of this book are highlighted in the form of *definitions* and *key propositions*.

Conventions on numberings: The numbering of figures follows the following pattern: “Fig. SecNo.-No.,” where “SecNo.” is the two-digit section number, and “No.” is the number of the figure within the section. Example: “2.1-2” is the second figure of [sec. 2.1](#). In the same way, *definitions* are numbered with “(Def. SecNo.-No.),” *key propositions* with “(Prop. SecNo.-No.),” and *examples* or other *highlighted* passages with “(SecNo.-No.).” Example: “(Def. 4.2-2)” is the second definition of [sec. 4.2](#); “([3.1-4](#))” is the fourth example of [sec. 3.1](#). *Quotations:* Double quotes are used for the purpose of literal quotation and single quotes for stylistic quotation (scare quotes).

At the end of every chapter, the reader finds hints on selected literature and exercises. This book concludes with a comprehensive bibliography and several indices.

This textbook is a revised and significantly extended English elaboration of my German Introduction to Philosophy of Science (“Einführung in die Wissenschaftstheorie”). The introduction has been developed from a series of lectures in the Philosophy of Science that I have given at the Universities of Salzburg, Erfurt, Düsseldorf and, less frequently, at other universities (including California at Irvine, Yale, Vienna, Helsinki, Graz, and Fribourg).

Parts of the text of the German edition have been pretranslated into English by Fergal Treanor. For valuable help concerning matters of content, I am indebted to Ioannis Votsis, Paul Thorn, Hannes Leitgeb, Wolfgang Spohn, Martin Carrier, Franz Huber, Markus Werning, Oliver Scholz, Andreas Hüttemann, Markus Schrenk, Hans Rott, Wolfgang Balzer, Gernot Kleiter, Theo Kuipers, Ilkka Niiniluoto, Erik Olsson, Stephan Hartmann, Jon Williamson, Igor Douven, Ken Gemes, James Ladyman, John Worrall, Stathis Psillos, James Hawthorne, Alan Musgrave, Mathias Frisch, Clark Glymour, Kevin Kelly, Horacio Arló-Costa, and Alvin Goldman. What remains on my side is to wish the readers of this book that it shall bring them much fun and intellectual profit.

Gerhard Schurz
Düsseldorf, March 2013

Foreword to the Second Edition

Since 2013, much has happened in the philosophy of science. Therefore, the Second Edition contains a lot of new content, concerning controversies about the role of values in science (sec. 2.10), results about meta-induction (sec. 2.11), debates concerning the question of scientific realism (sec. 5.8), and the mechanistic philosophy of science (sec. 6.4.1). In compensation, other and more technical parts have been shortened. Moreover, new literature has been added all over the place. The exercises with selected solutions have been extended. The index of subjects has been expanded so that it can be used as a *glossary* : for important terms, the page numbers on which the term is defined are shown in **bold**.

I would like to extend my thanks for valuable content-related help from the first edition to include Karel Lambert, Brian Skyrms, Branden Fitelson, Gila Sher, Johan von Benthem, Lorenzo Magnani, Holger Andreas, Mario Günther, Peter Brössel, Tom Sterkenburg, Leah Henderson, Richard Dawid, Matias Osta Vélez, Jochen Briesen, Gustavo Cevolani, Atocha Aliseda, Petr Spelda, Elke Brendel, Corina Strößner, Alexander Gebharder, Christian Feldbacher-Escamilla, Gottfried Vosgerau, Alexander Christian, Karim Baraghith, Niklas Parwez, Avril Styrman, and Christoph Schamberger.

Gerhard Schurz
Düsseldorf, September 2025

1 Introduction

Where Do We Stand?

DOI: [10.4324/9781003675075-1](https://doi.org/10.4324/9781003675075-1)

This chapter has been made available under a CC BY-SA license.

1.1 Tasks and Aims of Philosophy of Science

Science plays a determining role in the life of today's society. In what sense and to what degree can we trust the results of science? There is a scientific discipline, which examines such questions systematically: the *Philosophy of Science*, which in the German language is called "Wissenschaftstheorie" ("Theory of Science"). Following the German usage of the word, I always understand "science" ("Wissenschaft") in its *broad* meaning – including besides natural sciences, also social sciences, humanities, etc.

The discipline of philosophy (or theory) of science investigates how scientific knowledge works – its goals and its methods, its achievements and its limitations. Like any other scientific discipline, philosophy of science has historically developed out of philosophy. In the present days, it is practiced both by philosophers and by scientists of various disciplines.

One distinguishes between the *general* philosophy of science and philosophies of *particular* sciences. The latter ones are concerned with special kinds of disciplines such as philosophy of physics, biology, psychology, the social sciences, or the humanities. The general philosophy of science discovers those components of knowledge, which are more or less common to all scientific disciplines. Its main questions are:

- i. How is a scientific language constructed?
- ii. What are the rules of correct scientific reasoning or argument?
- iii. What are the distinctive features of a scientific observation or 'datum'?
- iv. What is a scientific law, and what does a scientific theory consist of?
- v. How are law hypotheses and theories empirically tested? When do they count as confirmed, and when as refuted?
- vi. What are the criteria for scientific progress?
- vii. What is achieved by a scientific explanation and what is causality?

This book focuses on these questions of the general philosophy of science, and the following chapters are arranged in corresponding order to the above questions. At the same time, we shall also deal with applications and topics of the philosophies of particular sciences, and examine characteristic *differences* between different branches of science. The most *general* questions of philosophy of science include:

- viii Is there an objectively discoverable reality and an objective truth?
- ix What connection exists between science and value judgments?

Question viii lies on the border between philosophy of science and epistemology (see [sec. 2.8](#)), and question ix on the border with meta-ethics (see [sec. 2.5](#), [2.10](#)).

Beyond its fundamental significance, which consists in providing answers to the above questions, philosophy of science also has a number of significant *applications*, both inside and outside science.

The *internal* applications of philosophy of science consist of the supply of foundational knowledge, which aids the scientist's decisions concerning new or controversial questions. Internal applications also consist of working out *interdisciplinary* common ground and providing *argumentative* competence in critical reflection. Not least, philosophy of science has often paved the way for new scientific disciplines (such as probability theory and statistics, measurement theory, or parts of cognitive science).

Among the *external* applications of philosophy of science to problems in a societal context, two in particular should be emphasized:

- a. The *demarcation* problem is highly significant in society. In this context, it consists of the question of which of our ideas have a claim to the status objective scientific knowledge that should be taught in public educational institutions, as opposed to subjective opinion, political values, ideologies, or religious convictions. This question became politically explosive in the controversy surrounding the creationism movement in the United States. For example, judge W. R. Overton in 1981 (McLean vs. Arkansas) and judge Jones in 2005 (Kitzmiller vs. Dover) based their decisions against the teaching of creationism in school on demarcation criteria between science and religious belief proposed by philosophers of science (cf. [Bird 1998](#), 2–9; [Branch 2011](#), 174). To tell the full truth, there were also philosophers of science, such as Laudan ([1983](#)), who have argued against the possibility of demarcating science from pseudoscience, and they have frequently been quoted by creationists as support for their position (to the embarrassment of other philosophers; cf. [Pennock 2011](#), 176ff).
- b. Equally significant is the critical role of philosophy of science concerning the commercial and political utilization of science, which counteracts the danger of *abusing* science and its results. Politicians, the media, and business leaders enjoy quoting expert opinions, which are presented all too often in a false or one-sided way in order to serve preordained purposes ([sec. 2.10.1](#), [4.3.6 – 7](#), [4.4](#)). But also applied scientists are often tempted to put subjective (e.g. financial) interests above objectivity. An example is the discovery of systematic biases in pharmaceutical research toward the financially supporting firms; philosophers of science have promoted scientific journal policies that may help to counteract such biases ([Brown 2008](#)).

1.2 On the Historical Development of Philosophy of Science

Although the term “philosophy of science” was only introduced in the 20th century, the philosophy of science is as old as science itself and had a parallel development (see [Losee 2001](#)). The history of the philosophy of science begins with *Aristotle* (384–322 B.C.), the ancient world's great systematizer of knowledge. Aristotle was much more oriented toward experience than was his teacher Plato. He was, nevertheless, like most philosophers after him, a follower of the *foundationalist* epistemological program. In this epistemological program, it is presumed that real knowledge is only possible if it stands on a foundation of certain and necessary principles, which are arrived at *not* through uncertain experience, but through rational intuition (cf. [Triplett 1990](#), sec. I; [Albert 1985](#), 18–28). Aristotle spoke in this context of “intuitive induction” ([Losee 2001](#), sec. Ia). In the current philosophy of science, by contrast, the *fallibilist* epistemological program has the upper hand, which concedes that our understanding of reality is basically *fallible*, and our scientific knowledge can be more or less well confirmed, but it cannot be guaranteed to be free of error. In a broader sense, of course, the scientific method is foundation-oriented, too, insofar it justifies its knowledge from most evident insights (primarily observation and logic) by inferences that are as reliable procedures possible – in contrast to coherentist epistemological programs, according to which the mutual coherence of beliefs is sufficient for their justification (cf. [Schurz 2024a](#), [sec. 2.1 – 2.2](#)).

Considerable advances in science and its philosophy had already been achieved by the late Middle Ages ([Crombie 1959](#)). Under the influence of philosophers such as *Roger Bacon* (1214–1292) or *William of Ockham* (1295–1349), the foundationalist program was gradually replaced by empirically oriented methodology. Of lasting importance is the principle of *Ockham's razor* , according to which one should reject all theoretical assumptions which are not needed to explain the observable. The decisive breakthrough, however, did not take place until the early modern period, the Philosophy of the Enlightenment. Building on the great successes of scientific method in this historical period – thanks to pioneers such as *Galileo Galilei* (1564–1642), *Isaac Newton* (1642–1727), and (later) *Charles Darwin* (1809–1882) – the influential *empiricist* movement was established, beginning with *Francis Bacon* (1561–1626) and *John Locke* (1632–1704) until *David Hume* (1711–1776) and *John Stuart Mill* (1806–1873). In the same period, the foundationalist epistemological program developed into the *rationalist* movement, prominently represented by *René Descartes* (1596–1650), *Gottfried W. Leibniz* (1646–1716), and *Immanuel Kant* (1724–1804).

The essential difference between the empiricist and rationalist philosophies can be characterized as follows. For the empiricists, sentences which may be justified *a priori* – that is, by pure reason and with rational certainty – are limited to so-called analytical sentences, that is, sentences whose truth is based on logic and conceptual convention. Such sentences possess no realistic content; they say nothing about the real world – only *synthetic* sentences may do this, but they can only be justified *a posteriori*, that is, based on experience. For the rationalists, on the other hand, *a priori* justifiable sentences with realistic content do exist, namely the so-called synthetic *a priori* sentences. However, both Descartes' and Kant's attempts to justify synthetic *a priori* principles of the sciences were unsuccessful and have been refuted by the further development of science. The development of the modern philosophy of science is closely bound with the insight that the foundationalist epistemological program was not tenable – at least not in its 'strong' version as explained above.

For a long time, classical systems of philosophical empiricism too had reductionist and foundationalist tendencies. The skeptical consequences of empiricism, which called for epistemic humility, were first recognized and radically formulated by Hume. Hume showed that the two central tenets of scientific method, the *causality* and the *induction principle* , are neither logically nor empirically justifiable. The causality principle says that all regular successions between events are produced by cause–effect relations; but Hume objected that all what one can observe are the successions of the events, while the idea that the earlier event or 'cause' produces the later event as its "effect" is metaphysical fiction (see [sec. 6.6.2](#)). The induction principle asserts that regular connections between events that have been observed in the past will also hold in the future; but Hume argued that this inference can neither be justified by logic nor by experience (see [sec. 2.11.1](#)). The problems involved in the justification of the principles of induction and causality have occupied philosophers up to the present day.

The apparent *insufficient justifiability* of empiricism put wind in the sails of subsequent rationalist movements, in which the foundationalist epistemological program, however, was gradually given up or replaced by pragmatized or historicized 'rationalisms' (see [sec. 1.3.2](#) –1.3.3). In the 20th century, *post-empiricist* and *post-rationalist* approaches have moved considerably closer to each other, and in the context of this tension, present-day philosophy of science has developed.

1.3 Philosophical Positions in Contemporary Philosophy of Science

1.3.1 Logical Empiricism

One of the most important origins of modern Philosophy of Science and Analytic Philosophy is logical empiricism, which was developed in particular by the Vienna Circle, but also by others, for example by members of the Berlin circle around Hans Reichenbach (1891–1953). The *Vienna Circle* was a group of philosophers and scientists in Vienna, whose core consisted of Moritz Schlick (1882–1936), Otto Neurath (1882–1945), and Rudolf Carnap (1891–1970). In the tradition of Ernst Mach (1838–1916), this group attempted to reestablish empiricism and scientific philosophy in general (cf. [Stadler 2001](#)). What was new about their situation was the development of *modern logic*, which had not been founded until the end of the 19th century. Modern logic made it possible to represent any language-based epistemological system with mathematical precision, and so the Vienna Circle hoped they had found at last the methodological equipment for a philosophy which could progress in a scientific manner (cf. [Schlick 1930/1931](#), 155f). In the middle of the 1930s, the Vienna Circle, with political leanings from liberal to socialist, had to disband under pressure from the National Socialists. The majority of its members emigrated, some via other countries, to the United States, where the logical-empirical movement, uniting with related Anglo-American directions, developed into Analytic Philosophy and gained a foothold in Europe in the postwar period.

Even today, many textbooks present logical empiricism as if it had represented a single dogmatic position; in reality, however, its proponents were highly capable of learning and continuously revised their positions. What today's philosophy of science can learn from logical empiricism is not so much any particular thesis but rather the high standards of conceptual and argumentative precision. In the phase up to 1935, the philosophy of science of logical empiricism was restricted by *positivism* and *reductionism*. In the later phase, the logical empiricists gradually rejected their restricted positions and replaced them with theses, to which the designations "empiricism" or "positivism" in the usual sense *no longer apply*. In one regard, logical empiricism has parted company with its classical predecessors early on: with the refutation of the classical empiricist thesis that at least 'direct' observation reports are infallible. In the so-called protocol sentence debate, Neurath and Carnap – but not Schlick – became convinced that even observation sentences such as "there is a table there" are in principle fallible. Empiricism, according to this new view, just meant that observation sentences have an epistemically favored role within the whole system of discovery (see [Carnap 1932/1933](#); [Neurath 1934](#), 113).

Empirical Reductionism, by contrast, was retained for quite some time. Classical empiricism advocated the following reductionist thesis: all 'serious' scientific concepts must be translatable via a chain of definitions into observational concepts. This program of reduction was adopted by logical empiricism in its early phase ([Carnap 1928a](#)). Later, however, the view prevailed that *theoretical concepts* such as "power," "electrical field," or "human character" *cannot* be defined by observational concepts, but rather go far beyond the immediately observable ([Hempel 1951](#); [Carnap 1956a](#)). In the so-called *standard* philosophy of science of the 1960s, the classical empiricist epistemological model had already broken down. What remained was a *minimal empiricism*, which consisted of the requirement that scientific theories must have empirical consequences, by means of which they can be tested.

A similarly modest empiricism is defended by contemporary empiricists such as van Fraassen (1980, [2002](#)). He does not only reject infallibilism and reductionism but also the *linguistic* orientation of logical empiricism. Instead, he defends a semantic conception of theories, in which theories are regarded not as sets of sentences, but as sets of conceptual models, similar as in structuralist philosophy of science (cf. [sec. 3.1.2](#)). Characteristic of van Fraassen's empiricism (1980, sec. 1) are two ingredients: first, a defense of the concept of theory-independent observation against Kuhn-followers (see [sec. 1.3.3](#)). And second, his rejection of scientific realism in favor of an instrumentalist position, according to which the theoretical (i.e. non-observable) part of a theory can only be reasonably said to be more or less *empirically adequate*, but not true or false. In this latter respect, van Fraassen is more restrictive than Carnap (e.g. 1928a) who always argued for ontological

pluralism , according to which both realism and instrumentalism are possible but epistemically undecidable practical attitudes toward scientific theories.

Two boundaries , which early logical empiricism had tried to maintain, became permeable in contemporary ‘minimal’ empiricism. *First* , the boundary between *science* and *metaphysics* became porous: it turned out that principles, which are of no empirical consequence *in isolation* , may give rise to new empirical consequences when joined with other theoretical sentences (see [sec. 5.4.2](#) ; [Hempel 1951](#) , 115; [Stegmüller 1969](#) ff, vol. II.1, 293–295). In this way, the *post-positivist* philosophy of science gained new access to the discussion of metaphysical principles such as realism and causality. Conversely, “inductive” branches of metaphysics emerged that sought to justify their principles not by pure thought but through empirical science (cf. [Hüttemann/Schurz 2025](#)). *Second* , the boundary between analytic and synthetic sentences became porous when it was recognized that the meaning of theoretical terms is not determined by isolated definitions, but by the respective background theory as a whole ([Carnap 1956a](#) , [1963](#)). In parallel fashion, Quine ([1951](#)) argued that in *natural* , empirically observed languages, no clear criterion exists by which one could undertake a demarcation between analytic and synthetic sentences.

As a consequence of this, analytic philosophy’s *linguistic* orientation (e.g. [Carnap 1950a](#)) was gradually given up. Linguistically oriented philosophers try to answer all philosophical questions by means of semantical concept analysis. But if a clear divide between (analytic) questions of meaning and (synthetic) questions of fact is missing, this exclusively linguistic orientation is no longer tenable. For this reason, Carnap ([1956b](#) , suppl. E) started to work on a system of pragmatics. Later on, Quine ([1974](#)) developed a *naturalistic* view, according to which science and philosophy form a continuum, the latter’s theses also having to be corroborated by experience.

1.3.2 Critical Rationalism

Also associated with the Vienna Circle was *Karl Popper* (1902–1994), the founder of so-called critical rationalism. From the beginning, Popper’s theory of science displayed the very non-reductionist tendencies that logical empiricism had accepted only after years in transition. According to Popper, scientific theories may go any distance beyond our experience, as long as they are empirically testable. Popper’s tests consist of *falsification attempts* , which should be as severe as possible. Popper supported his emphasis on falsification (instead of verification) by citing the *asymmetry* of *verification* and *falsification* of strictly universal hypotheses such as “all metals conduct electricity.” Such hypotheses cannot be verified by any finite number of observations, because they speak about a potentially infinite domain of instances – in our example, all metals in the past and in the future. However, these hypotheses can be falsified by a single observed exception ([Popper 1935](#) /2002, sec. I–IV).

For Popper, falsifiability was the decisive criterion of *demarcation* between science and speculation: scientific theories are not verifiable, but they must be falsifiable. Once falsified, they are *rejected* as false; but if they withstand manyfold falsification attempts, they are viewed as *corroborated* .

Popper’s falsificationism was also subjected to diverse criticism in the discussion which followed. Lakatos ([1970](#)) pointed out that scientific theories are almost never rejected on the basis of one single counterexample, but may be immunized against conflicts with recalcitrant data by means of modifications that are usually more or less *ad hoc* (see [sec. 5.6.1](#)). Popper’s *anti-inductionism* – that is, his thesis that science could progress without making any use of induction – also came in for harsh criticism (see [sec. 2.7.2](#)).

Popper’s critical rationalism is far removed from classical rationalism. Nowhere in Popper’s philosophy is it claimed that knowledge of reality can be grounded in non-empirical a priori intuition; on the contrary, Popper always rejected this. Thus, Popper’s philosophy is *post-rationalist* just as that of the late logical empiricists is *post-positivist* . Popper stressed more strongly than late logical

empiricism, however, that observation sentences are not only fallible, but also *theoretically biased* , so that the borderline between observational and theoretical terms cannot be sharply drawn (1935/2002, sec. V.30 addendum; new app. X). With this argumentation, critical rationalism opened its doors to the attacks from historical relativism, to which we turn now.

1.3.3 The Historical Account and Relativism

At the end of the 1950s, the historian of science Thomas Kuhn ([1962](#)) rattled the scientific world by presenting a fundamental attack on the standard philosophy of science. Factual science, said Kuhn, behaves differently from the normative ideals ‘prescribed’ by the philosophers of science. The alternative scientific model developed by Kuhn takes more of a historical-sociological than a logical-cognitive approach. According to Kuhn, scientific development takes place on the basis of so-called paradigms, such as the paradigm of classical (Newtonian) physics or (Darwin’s) theory of evolution. A Kuhnian paradigm contains at least the following three components: (i) very general theoretical principles or models, (ii) typical examples of successful applications, and (iii) methodological-normative assumptions (cf. [Kuhn 1977](#) ; [Hoyningen-Huene 1993](#) ; [Schurz 1998a](#) , [2014c](#)). The sociological correlate of a paradigm is a *scientific community* , a community of experts in the given field, who stick to a paradigm and work on its further development. According to Kuhn, a paradigm determines not only the fundamental principles, but also the major open problems and the *interpretation* of observed data – no, it even determines the observed data *themselves* , because, according to this view, *all observation is theory-laden* , and hence theoretically biased. For Kuhn, there is no observation which is independent of background theory or paradigm. Kuhn takes this strong thesis from Hanson ([1958](#)).

According to Kuhn, scientific development takes place in two successive phases, a *normal* and a *revolutionary* phase. Common acceptance of a paradigm makes the continuous growth of knowledge possible in the normal phase of a scientific discipline. When recalcitrant data, so-called anomalies, resist explanation by the given paradigm, the conflict is settled by more-or-less *ad hoc* modifications of current versions of theories within the accepted paradigm. Should such anomalies become more frequent, however, younger scholars will begin to look for a *new* paradigm. As soon as one is found, scientific development enters a revolutionary phase for a certain time, in which two paradigms struggle for supremacy. Kuhn’s examples in (1962) include the transition from Ptolemaic to Copernican astronomy and from Newtonian to Einsteinian physics. Kuhn claims that there are no common standards of rationality during a paradigm change and all empirical data are seen by the new paradigm in a new way. For this reason, the two competing paradigms are, according to Kuhn’s ‘notorious’ *incommensurability thesis* , rationally incomparable. So the struggle for supremacy does not take place in the form of a comparison of cognitive achievement, but rather in the form of a *power struggle* , in which supporters of the old paradigm eventually die out, as a result of which the new paradigm prevails, and a new normal scientific phase is heralded in.

The standard philosophy of science was unable to effectively counter the argument of the theory-dependence of scientific observations and its relativistic consequences. For all questions about the cognitive nature of observation had been banned by the narrowed perspective of the *logic* of science. For Popper, the question of what should be called an observation sentence was a question of pure convention (1935/2002, sec. V.30). The logical empiricists saw it in much the same way. It will be shown in [sec. 2.8.2](#) that a different picture results as soon as we take on an *empirical-cognitive* perspective. After all, if there were no difference between what experience *shows* , and what theory leads us to *expect* , then empirical science would be an exercise in *permanent self-deception* . Any theory of the empirical sciences which cannot explain this central difference must be extremely unsatisfying. Moreover, the loss of a viable criterion of empiricism would also undermine any credible distinction between real and fake facts. After right-wing movements began to criticize scientifically established facts, for example, concerning the climate crisis, as “leftist ideology,”

philosophers with left-leaning views tend to embrace once again the experience-based objectivity that was criticized by them some decades ago as an expression of “positivist ideology.”

Kuhn’s alternative model of science was the subject of much controversial discussion in the decades that followed. As well as an opposition to Kuhn, both a moderate and a more radical branch of Kuhn-followers was formed (cf. [Hoyningen-Huene 1993](#) , 212f). The more radical group gave relativism a strong impetus – most prominently in Feyerabend’s “Anarchist Epistemology” (1975), who with his maxim “anything goes” argued against methodological rules that are guiding or should guide science. The moderate Kuhnians, by contrast, aimed to separate Kuhn’s insights in the philosophy of science from the exaggerations of his “radical interpretation.” Kuhn himself later softened his incommensurability thesis and granted limited recognition to the possibility of rational comparisons between paradigms. ¹A moderate Kuhn-follower, and a Popper-follower at the same time, is Imre Lakatos and his conception of scientific research programs. Another moderate Kuhn-follower is Larry Laudan ([1981](#)), who argued that the empirical success of a scientific theory is not necessarily a reliable indicator of its truth. Such moderately Kuhnian developments have brought important insights to light for the philosophy of science (see [sec. 5.6](#) and 5.8). A further controversy centered around Kuhn’s thesis that in its normal scientific phases a scientific discipline is always dominated by a single paradigm. Scientists and philosophers of science have pointed out that, contrary to Kuhn’s thesis, disciplines outside of physics often exhibit a persistent coexistence of rival paradigms (cf. [Schurz/Weingartner 1998](#) ; [Kornmesser/Schurz 2014](#)).

1.3.4 Metaphysical and Scientific Realism

In reaction to the Kuhnian challenges of the logical-empirical model of science, many philosophers have suggested returning to the metaphysical approaches of realism and necessity, from which the philosophy of science in the early 20th century had tried to free itself. For example, it is a consequence of post-positivist philosophy of science that the term “mass” actually has different *meanings* in Newtonian and Einsteinian physics. This theory-dependence was also one of the main pillars of Kuhn’s incommensurability thesis. The early Putnam (e.g. 1975b, 219ff, 229–235, 240f), together with Kripke ([1972](#) , 55–59), developed the following argumentation against this consequence: scientific progress in the sense of a movement toward objective truth is only possible if the *reference* of scientific terms, that is, their relationship to reality, *is rigidly fixed* with metaphysical necessity, in the sense that this reference stays the same in all possible theories and across all possible worlds. Kripke and Putnam assumed the reference fixation to take place by acts of baptizing, together with a rigid sameness-relation between natural kinds, with the result that statements such as “water is H₂O” become true by “metaphysical necessity” (see also [Shoemaker 1980](#)). Fine ([1994](#)) also regards such identity relations as metaphysical necessities constituted by essential properties.

We agree with Kripke and Putnam that the meaning of natural kind terms such as “water” cannot be learned by an empirical definition (as in [Carnap’s 1956b](#) account), but by acts of ostension (“this is water”) grasping the right sameness-in-kind relation between water exemplars. But as a number of arguments have shown, scientific knowledge cannot determine sameness-in-kind relations and the reference of (theoretical) concepts with metaphysical necessity. ²Of course, classification systems for scientific (e.g. chemical or biological) kinds can be fixed *analytically* – that is, through definitional conventions (cf. [Sidelle 1989](#)). However, which classification system accurately reflects reality depends on the theoretical background knowledge ³which undergoes historical change ([Hendry 2008](#)). The idea that a necessary relationship between language and experience can be established independently of experience is in the end a regression toward a foundationalist epistemic model. In fact, the later Putnam ([1977](#) , [1982](#)) relinquished his earlier metaphysical realism by arguing as follows: if neither our empirical nor our theoretical knowledge can determine the reference of our concepts in a unique way, then *nothing* can do this – the consequence is an

unavoidable “indeterminacy of reference.” The same indeterminacy of reference has been worked out by Quine, but while Quine remained a naturalistic realist, the late Putnam adopted a more-or-less idealistic position that he called “internal realism.”

An account of necessity that is metaphysical in a weaker sense than the above ones is that of Armstrong ([1983](#)). According to him, laws of nature supervene on an assumed necessitation relation between “universals” (such as “water” and “H₂O”), which is, however, *contingent* – that is, these necessities may be different in different possible worlds ([sec. 6.5.6](#)). A similar middle ground between metaphysics and empiricism is Lewis ([1973](#) , [1986a](#)), who developed a possible worlds semantics for counterfactual and causal conditionals, but at the same time defends an elaborated Humean account to laws of nature, according to which laws supervene on the ‘mosaic’ of empirical facts ([sec. 6.6.5](#) ; cf. [Schrenk 2017](#) , [sec. 4.3](#) , [5.3](#)).

Like metaphysical realists, proponents of scientific realism also advocate for the convergence of science toward truth – but not on the basis of metaphysical postulates, but based on the empirical success of science. To support this position, scientific realists require robust arguments that allow them to infer the approximate truth of a theory from its empirical success. The most well-known argument of this kind is Putnam’s *no miracles* argument (1975a, 73). According to this argument, the empirical success of a theory would be a sheer miracle unless one assumes that the theory is at least approximately true. The no miracles argument has been subject to extensive criticism, notably by Laudan ([1981](#)), who, in his *pessimistic meta-induction* , argued that virtually all empirically successful theories of the past have eventually been shown to be false, whence it is inductively highly probable that our current best theories will suffer the same fate. Other philosophers of science, such as Boyd ([1984](#)) and Psillos ([1999](#)), have defended scientific realism against such anti-realist arguments, leading to a scientific realism debate that continues to this day (see [sec. 5.8.3](#) – [5.8.8](#)).

1.3.5 Naturalism and the Cognitive Turn

Modern naturalism goes back primarily to Quine ([1951](#) , [1960](#)), who rejected the analytic-synthetic distinction, thereby dissolving the clear boundary between scientific philosophy and empirical science. For Quine, naturalized epistemology becomes an empirical discipline, which is “scientifically investigating man’s acquisition of science” ([Quine 1974](#) , 3). The naturalistic program was emphatically taken up by a number of young epistemologists and philosophers of science (e.g. [Kitcher 1992](#) ; [Papineau 1993](#) ; [Kornblith 1994](#)).

Philosophical naturalism found expression in several philosophical programs. A first concretization of naturalism is *evolutionary epistemology* (e.g. [Campbell 1984](#)): many aspects of humans’ cognitive capacities and achievements can be explained by facts of evolution, in particular regarding perceptual knowledge. It has to be critically remarked, however, that naturalistic approaches themselves make certain epistemological assumptions which are *in need of justification* , for example concerning realism and induction, and so they cannot entirely replace classical epistemology. There is, however, an epistemological school of thought that indeed amounts to a *renunciation* of justification in this regard, namely epistemological externalism. According to its most well-known version ([Goldman 1986](#)), a belief counts as justified if it is produced by an “objectively reliable” method, without requiring anyone to be able to justify the reliability of the method in question. A justification of inductive prediction methods is not sought; instead, it is simply asserted that a predictive method is justified if it is reliable, that is, if its predictions will come out as true. Otherwise, it is considered unreliable. The externalist treatment of realism works in a similar way. However, since all nonscientific ‘methods’ – ranging from creationist doctrines of salvation to clairvoyance – also claim reliability for themselves, externalist epistemology fails to demarcate science from nonscience and thereby condemns itself to epistemic uselessness (see [Schurz 2024a](#) , [sec. 2.4](#)).

A second concretization of naturalism is the so-called cognitive turn. This turn resulted from the fact that meanwhile the cognitive sciences had attained the status of an interdisciplinary subject area in their own right, in which the basis of humans' cognitive and scientific capabilities is examined not only logically in philosophy but also empirically in psychology or linguistics and computationally in Artificial Intelligence research (cf. [Thagard 1996](#) ; [Anderson 2000](#)). As a result, several philosophers have started efforts to bring the approaches of the cognitive sciences into philosophy of science and vice versa (e.g. [Goldman 1986](#) ; [Bechtel 1988](#) ; [Bishop/Trout 2005](#)). The recent movement of "experimental philosophy" has even started to examine philosophical intuitions by empirical investigations instead of the traditional "armchair method," with interesting results concerning cultural differences ([Machery et al. 2004](#) ; [Knobe/Shau 2008](#) ; [Horvath/Koch 2020](#)).

The third specification of naturalism is, finally, naturalistic philosophy of science. Like Quine, it assumes a continuity between the methods of the empirical sciences and those of the philosophy of science. Questions of value are not excluded from science, but they are understood in an instrumental way – as questions of the optimal choice of means for given ends ([Laudan 1987](#)) – which closely aligns with the value-neutrality approach discussed in [sec. 2.5](#) . By contrast, the influence of externalism must be viewed critically: the philosophical aim of justifying fundamental scientific methods in contrast to pseudoscience is largely abandoned, and the reliability of scientific methods and results is simply presupposed ([Godfrey-Smith 2003](#) , 151). This runs counter to the aim of theoretical clarification of science and the exposure of the ideological misuse of scientific authority, as outlined in [sec. 1.1](#) .

1.3.6 Structuralist and Semantic View of Theories

In search of a better defense against the Kuhnian offensive, Stegmüller ([1976](#) ; [1977](#) , vol. II., sec. 6; 1979b) suggested replacing the "statement view" of the standard philosophy of science with a so-called non-statement view. Stegmüller's suggestion is based on the idea, traceable to Suppes ([1957](#) , 246ff) and Sneed ([1971](#)), that it is better to reconstruct scientific theories as systems of set-theoretical *models* than as systems of statements in a logical language. Such systems of models are also called *structures* : hence, the name "structuralist" philosophy of science. In 1976, [Stegmüller \(1976](#) , 105) envisions this step as a radical upheaval. Later, he is more modest and argues for the transition from logical to set-theoretical formalization methods for "pragmatic and psychological reasons" concerning the simplicity of formal reconstructions (1979b, 65). For set-theoretical models are also characterized by statements, namely by statements in the language of set theory. Therefore, there exist straightforward possibilities of *translation* between the theory representations in the statement view and the non-statement view (see [sec. 3.1.2](#) ; [Schurz 2014b](#)).

While the statement-versus-non-statement debate turned out to be a bit of a red herring, the structuralist philosophy of science can claim significant achievements in the *detailed reconstruction* of individual scientific theories and theory nets (cf. [Balzer et al. 1987](#) ; [Balzer/Moulines 1996](#)). Moreover, the structuralist philosophy of science has received support from Anglo-Saxon philosophers such as van Fraassen (1980, sec. 2; 1989, sec. 9) and Giere ([1988](#) , [2004](#)), who called this position the "semantic view of theories" and developed it in a less formal way (cf. [Suppe 2000](#)). The semantic view of theories essentially consists in the claim that theories should be understood as models – a view later taken up by the "new mechanists" in the philosophy of science (e.g. [Glennan 2017](#) ; [Glennan et al. 2022](#) ; see [sec. 6.4.1](#)). However, the concept of a model remains vague: virtually anything can serve as a 'model' – from images to mathematical equations – as long as it serves to represent some aspect of reality (see [Glennan 2017](#) , 60; for criticism, see Chakravartty 2001; [Schurz 2014b](#)).

1.3.7 Pragmatic Philosophy of Science

In the 1980s, several authors spoke of a *pragmatic turn* in philosophy of science ([Stegmüller 1969](#) ff, vol. I, sec. XI; [Böhler et al. 1986](#) ; [Bernstein 2001](#)). Unfortunately, however, pragmatic philosophy and the use of the term “pragmatics,” in general, fall into two *opposing* camps. While one camp (e.g. Ch. S. [Peirce 1931](#) ff, van Fraassen 1980; [Rescher 1998](#) , [2001](#) ; [Schurz 1998](#) b) understands pragmatics in the *epistemically internal* sense, the other camp (e.g. [James 1907](#) ; [Rorty 1982](#) ; [Stich 1990](#) ; partly [Putnam 1995](#) ; Bernstein 2010) sees pragmatics in an *epistemically external* sense.

For the epistemically external camp, the pragmatic components of discovery are essential, but have nothing to do with its truth. By contrast, for the epistemically internal camp, it is these very pragmatic components, which are supposed to establish an objective conception of truth in a noncircular sense. The epistemically external camp makes knowledge (or justified belief) dependent on *any* kind of interests, be it predictive accuracy, political control, or emotional well-being, and sees therein a compelling argument for the unavoidable subjectivity of knowledge. In this sense, Stich ([1990](#) , 131ff) has argued that there are so many different conceptions of truth as there are different cognition-guiding interests. By contrast, for the internal camp, the reference to epistemically *internal* purposes is the very thing which makes possible a noncircular foundation of knowledge. For example, the reference to the cognitive purpose of maximizing true predictions enables a justification of induction that escapes Hume’s skeptical challenge ([sec. 2.11.2](#)). In conclusion, the external and the internal camp *differ so greatly* in their understanding of pragmatics, that they should not be subsumed under a single term.

1.3.8 Hermeneutics and Critical Theory

The *Philosophy of the Humanities* , which originated in the 19th century, goes back to Friedrich Schleiermacher (1768–1834), Wilhelm Dilthey (1833–1911), and Wilhelm Windelband (1848–1915). It started as a movement for the methodological demarcation of the humanities against the expanding natural sciences. Central to this philosophy is the use of *hermeneutics* , the study of interpersonal *understanding* (*Verstehen*) , for the purpose of such a methodological demarcation. The *methodological dualism* of the philosophers of the humanities has been brought to a head with the following thesis: in the natural sciences, we *explain* , but in the humanities, we *understand* . While scientific explanations are based on general law hypotheses, it was claimed that understanding in the humanities needs no recourse to such law hypotheses, because the human mind is not subject to strict law mechanisms, and understanding has direct access to the human mind.

Hermeneutics is part of a tradition from well before Schleiermacher, in which there was no mention of any rigid separation of the humanities from logic and natural sciences (see [Scholz 2001](#) , part I.B). The methodological dualism originated in the 19th century and had a strong influence on methodology debates of the 20th century (cf. [Von Wright 1971](#) ; [Manninen/Tuomela 1976](#) ; [Apel 1984](#)). The contemporary controversy in the humanities and social sciences about *quantitative* versus *qualitative* methods is partially a revivification of the old methodological dispute (cf. [Lamnek 1988](#) ; [Denzin/Lincoln 2000](#)). In response, the analytic philosophy of science has produced a number of approaches, in which the unity of hermeneutics and empirical methods is worked out (see [sec. 6.4.2](#) ; cf. [Hirsch 1973](#) ; [Follesdal 1979](#) ; [Taylor 1980](#) ; [Bühler 2003](#)).

Critical theory was developed by the philosophers of the so-called Frankfurt school, mainly Horkheimer, Adorno, Marcuse, and Habermas. It combines the methodologically dualist version of hermeneutics with the doctrine of the unavoidable interest-dependence of discovery and knowledge (see [Habermas 1966](#) ; [Apel 1984](#)). Critical theory also influenced the philosophy of science debate on the value-freedom versus value-ladenness of scientific research. We turn to this question in [sec. 2.5](#) and 2.10, where a modernized version of scientific value neutrality is developed – one that is compatible with value commitment in science and capable of withstanding relevant objections.

1.3.9 Radical Constructivism, Postmodernism, and the Science Wars

Radical constructivism was developed in the 20th century by biologists such as Maturana and Varela ([1992](#)) and psychologists such as von Glasersfeld ([1995](#)); philosophically, however, it is much older. The central argument of constructivism infers that, because our perceptions and descriptions of reality are constructed, there is no subject-independent reality – instead, *reality itself* is said to be a human construction. Even the early Kuhn embraced a constructivist position, claiming that “the world itself” has changed following a scientific paradigm shift ([Kuhn 1962](#) , chapter 10; see [Godfrey-Smith 2003](#) , 96).

In [sec. 2.8.2](#) , we will show that the argumentation of radical constructivism is based on a *fallacy* . Regardless of this, radical constructivism, in combination with the hermeneutic methodological dualism, has exerted its influence in the cultural and social sciences. There, it shaped French poststructuralism and the philosophy of postmodernism ([Lyotard 1984](#)), which reject scientific reality and its impact on human culture, instead promoting the idea that reality is entirely socially constructed. In the United States, these currents gave rise to the field of “Science Studies” and encouraged a left-leaning politicized understanding of science, as exemplified by feminist philosophy of science ([Harding 1986](#) ; cf. [Godfrey-Smith 2003](#) , sec. 9.1–9.4). Most pointedly, the social constructivist position can be found in the “science studies” of Latour and Woolgar ([1986](#)) and Knorr-Cetina ([1981](#)), where scientific laboratories (e.g. in nuclear physics or biochemistry) are studied by ethnographical methods of studying “foreign cultures” without engaging with the actual epistemic content being investigated, in order to learn something about the “social construction of scientific reality.”

The disregard of methodological standards and the misuse of scientific terminology met with strong opposition from the scientific community and led, in 1996, to a “counterattack” in the form of the *Sokal Affair* . In order to expose the lack of qualifications behind the wordplay of postmodern authors, Sokal ([1996](#)) published a hoax article titled “Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity,” which consisted of a jumble of quotes from postmodern thinkers. Despite its nonsensical content, the article was accepted and published by the postmodern-oriented American journal *Social Text* . When Sokal later revealed in another journal that his article had been a parody of postmodern discourse, the scandal was complete (cf. [Godfrey-Smith 2003](#) , 146f). The subsequently published book by Sokal and Bricmont ([1998](#)), entitled *Fashionable Nonsense: How the Postmodernists Abuse Science* , caused a major stir and became a catalyst for the so-called Science Wars ([Ross 1996](#)).

1.4 Introductory Literature and Exercises

1.4.1 Introductory Literature

A historically oriented introduction into the philosophy of science of the 20th and 21st centuries is Godfrey-Smith ([2003](#)). Ladyman’s introduction (2002) is focused on natural science and Rosenberg ([2008](#)) introduces to the philosophy of social science. Okasha ([2016](#)) gives a systematic but very brief introduction. An excellent introduction to the history of the philosophy of science is offered by Losee ([2001](#)). Stadler ([2001](#)) presents an extensive study of the Vienna Circle and its logical empiricism. For Popper’s philosophy of science, see [Keuth \(1998\)](#) , ed.). Hoyningen-Huene ([1993](#)) offers an introduction into the philosophy of Kuhn. [Hüttemann \(2021\)](#) offers an introduction into the metaphysics of science; for critical discussion of metaphysical accounts, see Ladyman and Ross ([2007](#)). On evolutionary epistemology, see Campbell ([1984](#)). The history and current situation of hermeneutics are dealt with by Bruns ([1992](#)).

1.4.2 Questions and Exercises

Sec. 1.1–1.2

1. What are the seven main questions in the general philosophy of science?
2. What is the major difference between the empiricist and rationalist ways of thinking in epistemology?

Sec. 1.3

1. What is understood by empirical reductionism?
2. What criterion of demarcation is suggested by critical rationalism?
3. What is a paradigm for Kuhn? Discuss Kuhn's incommensurability thesis.
4. How does scientific instrumentalism differ from scientific realism?
5. Discuss the methodological dualism of the philosophers of the humanities from a contemporary viewpoint.

Notes

1. Kuhn ([2000](#) , 33–57); cf. Okasha ([2016](#) , 87); [Seselja/Straßer \(2013](#) , 14f). Empirical data show that history of science proceeds much more cumulative than Kuhn suggests (see [Krauss 2025](#)). ↩
2. See [sec. 5.8.2](#) . Cf. Mellor ([1980](#) , 118f); Bird ([1998](#) , 108–120); Ladyman/Ross ([2007](#) , 290ff). ↩
3. In philosophy of science, “background knowledge” is not understood in the strict epistemological sense: it contains scientific beliefs that are rationally accepted but not necessarily true. ↩

2 The Question of Unity

In Search for Common Foundations of the Sciences

DOI: [10.4324/9781003675075-2](https://doi.org/10.4324/9781003675075-2)

This chapter has been made available under a CC BY-SA license.

2.1 Normative or Descriptive? The Method of Rational Reconstruction

Prima facie, we can identify two opposing views of the goals and methods of the philosophy of science:

1. According to the *normative* view, it is the goal of philosophy of science to say what science *should* be, and how it should be pursued. For this purpose, philosophy of science must determine in what scientific rationality consists, and on the basis of which criteria a scientific hypothesis may be justified.
2. According to the *descriptive* view, on the other hand, it is the task of philosophy of science to say what science de facto *is*, and how it proceeds. For this purpose, philosophy of science must describe and explain the existing sciences in terms of their historical development and their current structure.

The normative view is historically older, and currently still enjoys a wide following. Both logical empiricists and critical rationalists held this view. It was not until Kuhn ([1962](#)) set off the historical turn in philosophy of science that the descriptive counterpoint became popular. It can be found in the work of historically oriented philosophers of science (cf. [McMullin 1970](#) ; [Laudan 1977](#) , 162), of the structuralists (e.g. [Stegmüller 1976](#) , 267ff; [Balzer et al. 1987](#) , xvff), and of the naturalists (cf. [Giere 1999](#) , 160–163; [Ladyman 2002](#) , 4). So which position is right?

The position of the normativists can be traced back to Hans Reichenbach's ([1938](#) , 6f) distinction between the *context of discovery* , that is, the factual *genesis* of scientific knowledge, and its *context of justification* . The plausible prima facie argumentation of the normativists may be broken down as follows (e.g. [Popper 1935](#) /2002, [sec. 1.2](#)):

(2.1-1) *Normativist argumentation* :

Premise 1 : To assess the epistemic value of a hypothesis, *only* the question of rational justifiability of its truth claim is relevant.

Premise 2 : Questions about the hypothesis's historical development, the causes for its discovery, etc., are irrelevant to this purpose.

Conclusion : Therefore, philosophy of science should only examine science's rational justification and not its context of genesis – the latter being the remit of the history, sociology, or psychology of science, but having nothing to do with its philosophy.

Since Kuhn ([1962](#)), the argumentation of the normativists has been criticized in a very *sweeping* way. Here, we will attempt to clarify what is correct and what is flawed in this line of argument. The argument is certainly logically correct – assuming we take it to be analytically true that the philosophy of science is concerned with the assessment of the epistemic value of hypotheses. That is to say, *if* the premises of the argument were really true, then the conclusion would also be true. There is also *no error* in premise 1 – this thesis also seems to be analytically true, to the extent that science is concerned with the discovery of truth, and nothing else. But premise 2 is already infected by the 'germ' of the normativist fallacy. Because the

historical genesis of a scientific hypothesis includes not only so-called external factors which are epistemically irrelevant, but also internal factors which are epistemically relevant: for example, the experimental methods, which scientists de facto use, and the rational hypothesis evaluation, which they de facto apply.

The organic chemist Kekulé, for example, reported that his groundbreaking hypothesis of ring-shaped benzene molecules first appeared to him in a *dream* (see [Hempel 1966](#) , 16). This fact was, of course, epistemically irrelevant and played no role in the justification of the hypothesis. By contrast, the *Boyle-Mariotte law of gases* was discovered by means of the systematic extrapolation of data ([Langley et al. 1987](#) , 81) – and this fact played a decisive role in the justification of the hypothesis.

Philosophers should *learn* from the de facto methods and arguments of science as far as they are successful, and they can only do this by observing them. Instead of premise 2, all the normativists usually want to say, or at least all they should say, is the following *almost trivial* weakening:

(2.1–2) *Premise 2** : The epistemically *external* part of the history of the development of a hypothesis is irrelevant to the assessment of its epistemic value.

Of course, the conclusion of argument ([2.1-1](#)) no longer follows if premise 2 is replaced by the weaker premise 2*. And with this conclusion, the error of the normativist argumentation becomes *grave* . To realize this, we must simply ask the following question: how can philosophers of science *acquire* the appropriate rules and criteria for the scientific method? It would be naïve to presume that these rules and criteria might be gained through pure logic or mere intuition; science is much too complex for that. That would be like somebody trying to write a bridge-building manual without having examined the de facto practice of bridge building and typical examples of existing bridges. The philosophy of science, then, *can* and *should* learn from the real examples provided by the history of science. This brings us to the following result: although it is the *primary* task of the philosophy of science to explain the methods and criteria of *rational* science, it can only fulfill this task by *also* studying the *factual sciences* in their historical development.

Kuhn's criticism crystallized into the insight that the scientific models provided by logical empiricism and critical rationalism were too *simplicistic* to give a faithful picture of real science (cf. [Stegmüller 1979b](#) , 5; 1969ff, vol. II.3, 20). Reacting to this, young theorists proclaimed that the philosophy of science should even be restricted to the descriptive analysis of the sciences. But this position too is overstated. The question of the definition and criteria of scientific rationality must, of course, remain central to philosophy of science, as it is this question which defines philosophy of science as a discipline and sets it apart from the history and sociology of science. In summary, the philosophy of science is a discipline, which contains both descriptive and normative elements – an insight that has been worked out by [Stegmüller \(1969](#) ff, vol. IV.1, 9ff).

The philosophy of science's method may best be described as a *rational reconstruction* . My understanding of this method is different from Lakatos' understanding of it, which has been sharply criticized by Laudan ([1977](#) , 167–170). For Lakatos ([1971](#) , 119f), 'history' as understood by rational reconstruction is not the *actual* history, but a projection of ideals of rationality into history, and a selection of those parts of history that conform to that ideal. By contrast, in my understanding, the actual history forms one pole of the method.

More precisely, rational reconstruction moves between two poles: a (so-called) *descriptive corrective* , which should be adequately reconstructed, and a *normative corrective* , which contains rational norms. The descriptive corrective does not contain every piece of factual sciences. It contains only typical examples of de facto successful (i.e. unanimously and sustainably accepted) scientific discoveries. Moreover, it also contains typical 'counterexamples,' that is, examples of de facto refuted scientific hypotheses. Philosophy of science bases its reconstructions on these consensually accepted examples of scientific success or failure, though only in a preliminary and defeasible way.

The normative corrective consists of certain epistemic intuitions that are normally assumed by scientists, in spite of their various cognitive differences. That such a common ground exists is the position that will be defended in this chapter. First of all, the normative corrective contains one supreme epistemic goal or norm, which may be explicated as follows (cf. [Weingartner 1978](#) , [sec. 3.2](#)):

(Prop. 2.1-1) The *supreme epistemic goal* (G) of science is to find *true* and *content-rich* statements, laws, or theories, relating to a given domain of phenomena.

A sentence is the richer in content, the more *consequences* it has (see [sec. 3.7](#)). The restriction of the goal of science to true *and* content-rich statements is significant, because the probability and the content of hypotheses are often *inversely proportional*. We can maximize the truth chances of hypotheses by sticking to trivial tautologies like: “The Sun revolves around the Earth, unless it doesn’t.” Inversely, it is easy to present very substantial and impressive hypotheses if we disregard their truth value, for example, “I have invented a perpetual motion machine.” The true art of the scientist is to formulate hypotheses which have both empirical truth and substantial content and consequences.

In the above explication of the supreme goal of science, there appears the little word “truth.” One might object that the goal of truth is not metaphysically neutral. However, we understand this concept in a liberal way, leaving room for different metaphysical interpretations (see [sec. 2.8.1](#)). On the other hand, we don’t want to be too liberal about the meaning of “truth.” For people have rather different ideas of what truth is. While one person seeks truth in the platonic heaven, another person may equate truth with empirical success, and for a third one truth is what makes herself happy. To ensure that the goal of scientific discovery as explicated in [prop. 2.1-1](#) is not entirely vague concerning the meaning of “truth,” we need a *minimal* explication of what “search for content-rich truths” is intended to mean. This is provided by the second ‘normative’ component that is normally assumed by scientists as well as by philosophers of science, namely a minimal common epistemological model of scientific discovery. In [sec. 2.2](#), we will explain this minimal epistemological model of the sciences in detail. The key points are given here: it contains (1) a *minimal realism*, (2) *fallibilism* and critical attitude, (3) the struggle for *objectivity*, (4) a *minimal empiricism*, and (5) the struggle for *logical clarity*. Although this minimal epistemological model consists primarily not of normative, but of *descriptive* claims, we still regard it as part of the normative corrective of philosophy of science, as it gives clear meaning to the supreme goal of science, and is usually seen as a precondition of science.

Given the above described normative and descriptive correctives, the method of rational reconstruction may be precisely characterized as follows:

(Def. 2.1-1) *Method of rational reconstruction* : Develop generalized and logically precise *models* of scientific discovery, which on the one hand *fit the descriptive corrective*, that is, provide optimal explanations for the success of the examples and the failure of the counterexamples from the history of science, and which on the other hand may be *justified according to the normative corrective*, as providing optimal *means* of reaching the supreme goal of scientific discovery, given the assumptions of the minimal epistemological model.

In short, the philosophy of science abstracts generalized models from the descriptive corrective, which at the same time are (or should be) in line with the normative corrective. This is summarized in the schematic diagram in [fig. 2.1-1](#).

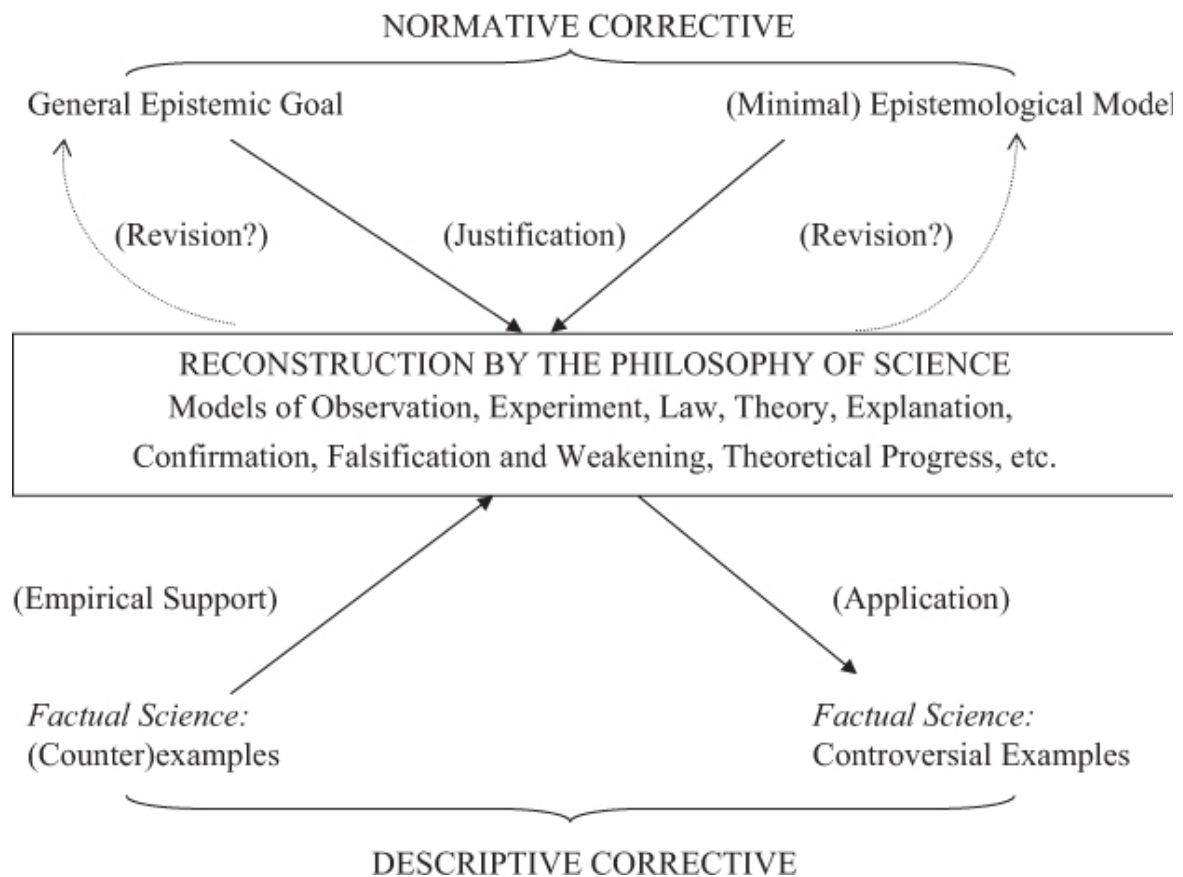


Figure 2.1-1 The method of rational reconstruction. [↗](#)

As long as both aspects of rational reconstruction – descriptive adequacy and normative justifiability – can be made to work in harmony with each other, the work of scientific reconstruction can be counted as successful and its results as *corroborated*. Of course, we cannot expect that everything we find in *factual science* will also be *rational* according to the normative corrective. For this reason, the descriptive corrective of the philosophy of science contains only *accepted* (counter) examples. When models developed by the philosophy of science have been corroborated, we can take the next step of applying them to *controversial* areas of scientific disciplines, in which no consensus has been reached – be it the objectivity problem of quantum mechanics, the teleology problem in biology, or the explanation-understanding problem in the humanities. It is exactly here that philosophy of science can supply valuable decision-making aids to the individual sciences. So, on the one hand, the philosophy of science learns from the individual sciences and is corrected by them; on the other hand, its reconstruction work can be instructive to the individual sciences and thus help them solve specific problems.

The goal of rational reconstruction, though, bringing together descriptive analysis and normative justification, does not always have to succeed. This applies, for example, when philosophers of science proceed from unrealistically high or one-sided normative standards, which will persistently disallow them to reconcile their normative corrective with the descriptive analysis of science. In such cases, the philosophy of science must go on to *revise* its own normative corrective. When it does this – this has already happened several times in the history of the philosophy of science – then the philosophy of science, to borrow Kuhn's terminology, is no longer in a normal scientific phase, but in a revolutionary phase.

To conclude this section, we will compare the model of rational reconstruction with two similar characterizations of the method of philosophy of science. Carnap ([1950b](#) , sec. 2) described the procedure of the philosophy of science as *explication*. Here, a vague concept of everyday or scientific language (the explicandum), for example, the idea of probability, is replaced by a precisely defined term (the explicatum), whereby the explicatum must differ from the explicandum insofar as it eliminates vagueness and ambiguity,

but it should also remain as close to it as it can and be as fruitful and simple as possible. Brun ([2016](#)) describes Carnap's method of explication as "conceptual re-engineering." Carnap does not explain, however, how exactly these partially contradictory demands should be combined. In comparison, the proposed model of rational reconstruction contains clear normative and descriptive guidelines; and by distinguishing normal and revolutionary phases in the philosophy of science, it takes into account the possibility that incoherencies can arise between the two. There are also some similarities between rational reconstruction and the method of reflective equilibrium according to Goodman ([1955](#) , 85–89) and Rawls ([1971](#) , 20f, 48–52). However, while the method of reflective equilibrium works purely *coherently* , aiming at the reciprocal adaptation of methodological rules and intuition, rational reconstruction is *empirically grounded* by its descriptive corrective. Moreover, the normative corrective consists not of mere intuition, but of epistemological *hypotheses* , together with a supreme epistemological norm.

2.2 Common Epistemological Assumptions

The minimal epistemological model, common to more or less all scientific disciplines, contains five epistemological assumptions (E1–E5):

E1 – Minimal realism : According to this assumption, there is a reality, which exists independently of the (given) epistemological subject. It is not assumed that all properties of this reality are knowable. The possibility of fundamental limitations to scientific knowledge is left open and cannot be answered a priori, but only based on the factual epistemic success of science. Scientific disciplines aim to provide statements which are as true and as content-rich as possible about clearly defined areas of this reality. The term "truth" is used here in the sense of the *structural correspondence theory* , according to which the truth of a sentence consists in its structural agreement with the part of reality described by it. This notion of structural correspondence, which was first precisely defined by Alfred Tarski ([1936a](#)), does *not* imply any intrinsic relationship or mirroring between language and reality. It also allows that theories which are strictly speaking *false* may be true in an *approximative* sense, or *close* to the truth ([sec. 5.6.2](#) , 5.11.1).

E2 – Fallibilism and critical attitude : There is no infallible 'golden path' to the truth in the sense of the correspondence theory. According to the assumption of fallibilism, every scientific statement is more or less fallible; so, we can never be *absolutely* sure of their truth. Fallibility goes hand in hand with a critical attitude, according to which no statement may ever be once and for all exempted from criticism. However, we can regard the truth of a statement as more or less probable, and in the limiting case as "practically certain," in which case we consider its error probability negligible. For this reason, everything depends on using empirical tests to find out more about the *probability* of a scientific hypothesis, in the sense of assumption E4 below.

E3 – Objectivity and intersubjectivity : According to this assumption, a statement's truth must hold objectively, that is, independently of the beliefs and value-attitudes of the epistemological subject, because by assumption E1, reality has subject-independent existence, and truth is the correspondence between statement and reality. So, the assumption of objectivity already follows from assumption E1. However, the characterization of objectivity as subject-independence does not help us in the practice of science, as only subjects make statements and formulate hypotheses. We have no direct access to objective truth; we can only approach it indirectly via – not certain, but probable – *criteria* .

The second part of assumption E3 says that a central criterion for objectivity and also for truth lies in the *intersubjectivity* of scientific statements. Intersubjectivity, correctly understood, is a plausible consequence as well as a plausible indicator of objectivity and truth. It is a plausible consequence because if a statement's truth can be established at all convincingly, then – at least in principle – it must be possible to convince any person of the statement's truth, provided the person is sufficiently competent, that is, cognitively trained and informed. And it is a plausible indicator, because the agreement of many mutually *independent* and competent (i.e. fallible but truth-biased) researchers on the hypothesis provably increases its truth chances. This is an instance of the *Condorcet jury theorem* (cf. [Bovens/Hartmann 2003](#) , [sec. 5.2](#) ; [Dietrich/Spiekermann 2023](#)).

E4 – Minimal empiricism : With the exception of formal sciences such as mathematics, which we will discuss later, the objects examined by any science must in principle be accessible to experience and observation. After all, reliable information about reality can only be gained through perceptual observation; for perception is our only form of direct informational contact with reality. Empirical observations, then, are a decisive referee in the scientific search for truth: scientific law hypotheses and theories must be *tested* in terms of them. This is the central statement of assumption E4. Observations are not infallible (that would contradict assumption E2), but they offer the easiest and fastest route to intersubjectivity and practical certainty. Observations are “epistemically privileged,” as one also says. The empiricism contained in assumption E4 is *minimal*, because different from empirical reductionism it is not claimed that all scientific statements must be traceable back to observations via a chain of definitions, or must be provable by observations. Scientific theories can and should contain theoretical concepts, by which they can talk about things which lie beyond what is observable to human senses. The decisive requirement of E4 is merely that statements about the unobservable must have *observable consequences* by which they may be tested.

E5 – Logic in the wider sense : Using precise logical methods to introduce terms, to formulate sentences, and to construct correct arguments, is the most effective way to approach the goal of content-rich truth (according to assumptions E1–E4). For the meaning of a sentence cannot be established until the terms used in it have been clearly defined. And only sentences with a precisely formulated meaning may be examined for their logical consequences. Finally, only after the consequences of a hypothesis are known, may this hypothesis be precisely empirically tested in accordance with assumption E4. In conclusion, the procedure of empirical testing requires the use of exact logical methods at all times – in the wider sense of logic, not restricted to deductive logic.

Our minimal epistemological model makes as few assumptions as possible, but as many as necessary, in order to explain the possibility of objective science. Some of its assumptions are worthy of further explanation. We call the realism in assumption E1 “minimal” because it only asserts the objective *existence* of reality. Whether and to what extent reality is objectively discoverable cannot, according to assumption E2, be answered a priori, but only through the success of scientific discovery methods. This sets our minimal realism apart from metaphysical realism, in all its varieties. Moreover, we don’t claim that science is impossible without the assumption of a subject-independent reality, but only that this assumption is *normally* made in scientific disciplines (see [sec. 2.8.1](#)).

To avoid misunderstandings, we should add that the correspondence theory of truth, while well suited as a *definition* of truth, does not offer any *criteria* to evaluate the truth of a sentence (cf. [Rescher 1973](#) , ch. 1; [Kirkham 1992](#) , 195). The path to such criteria is illuminated by assumptions E5 and E4: investigate the empirical consequences of your hypothesis and test them using empirical observations. For example, to find out the truth of the sentence: “this flower is red,” the correspondence-theoretical insight that this sentence is true if the flower is red will not get me anywhere; rather, I must be in a position to apprehend the object denoted by the ostensive individual term “this flower” and the perceptual quality expressed by the predicate “red” visually, in order to test the truth of the sentence by a simple act of perception.

In assumptions E4 and E5, we claim the existence of only two minimal levels, in which discovery finds its origin and its touchstone: observation, which allows us fragmentary but reliable and intersubjective access to reality, and logic in its wider sense, which provides a universal cognitive tool, but in itself says nothing about reality.

In assumption E3, we limited our characterization of *intersubjectivity* to “sufficiently competent and informed” persons. This limitation is necessary, because persons who do not *understand* the content of a hypothesis, or who do not possess the relevant *evidence* , are unable to test a hypothesis for its truth. This difficulty shows us that intersubjectivity can on no account be used as a *definition* of objectivity or truth. The latter suggestion has been made in the constructivist school of Kamlah and Lorenzen ([1984](#) , 119), and one century earlier by the founder of the philosophical school of *pragmati(ci)sm* , Charles S. Peirce ([1878b](#)). But competence is an uncertain and gradual criterion, meaning that any number of persons judged to be competent may still err collectively. Peirce attempted to save his consensus definition by equating truth with the consensus that an *ideal* community of researchers with an infinite amount of time would approach in the limit. But even in this counterfactual version, an intersubjective consensus does not have to agree with the

truth of correspondence theory. There are countless questions on whose true answer even an ideal community of researchers can never achieve consensus, because there is no evidential access to them, although there is no reason to assume that these propositions have an indefinite truth value – for example, “How many microbes were located in Caesar’s nose as he crossed the Rubicon?” In conclusion, intersubjectivity can at most be an uncertain criterion for objectivity.

2.3 Common Methodological Features

From the supreme epistemic goal (G) and the five epistemological assumptions (E1–5), there follow very plausibly four methodological features (M1–4), which are common to all empirical sciences in the broad sense. First, a word on terminology. By an *observation sentence*, we understand a sentence that expresses an *observable* state of affairs. We distinguish between *actual* and *potential* observation sentences. We call an observation sentence actual if the state of affairs expressed by it has indeed *been observed*. By contrast, an observation sentence is merely potential, if the state of affairs expressed by it has not been observed, but could be observed, for example, at a later time or in a different place. Merely potential observation sentences express empirical *predictions*, for example, “there will be a solar eclipse tomorrow.” We can now formulate the four common methodological features as follows (similarly in [Weingartner 1978](#), [sec. 3.1](#) – [3.7](#)):

- M1* : Science searches for hypotheses which are as general and as content-rich as possible and recorded in a scientific language. In all disciplines, these hypotheses include *laws* and *theories*, and in some (e.g. the historical sciences), they also include *hypothetical singular sentences*.
- M2* : Science searches for *actual observation sentences*, as many as possible (and as relevant as possible), which reflect the results of observations, experiments, or measurements.
- M3* : Science attempts, with the help of general and hypothetical sentences, to *explain* the currently known actual observation sentences, and to *predict* new, and as yet unknown, *potential* observation sentences.
- M4* : Science attempts to *test empirically* its general and hypothetical sentences by comparing the predicted (potential) observation sentences with the currently known (actual) observation sentences. If the latter are in agreement with the former, the prediction was successful (it then becomes a successful explanation) and the hypothetical sentence (law or theory) is *confirmed* or *corroborated*. If there is a contradiction between the latter and the former, the prediction was without success and the law or theory is *falsified*, or in the case of a merely statistical prediction, *weakened*.

We only speak of an *explanation* if the state of affairs to be explained has already come to pass and is known, while in a *prediction*, the predicted state of affairs is new and unknown. Viewed logically, both explanations and predictions take the form of deductive or probabilistic arguments (see [sec. 6.1](#) – [6.4](#)).

These four methodological features or modules are closely connected with the epistemological assumptions of science. Feature M1 results from the epistemic goal G, because laws and theories are the typical examples of general, content-rich sentences. M2 results from assumption E4 (sharpened by assumptions E2 and E3): to discover more about the conjectured truth of its hypotheses, science must collect empirical data. In order for general hypotheses to be tested on the basis of these data, the logical or probabilistic consequences of the hypotheses, in the form of predictions or potential explanations, must be investigated, and these consequences compared with the data – in this way, the methodological building blocks M3 and M4 are motivated by the assumptions E4 and E5.

Successful predictions and explanations not only have a bearing on the ultimate scientific goal of finding true and content-rich theories, but they are also of crucial importance for human *planning behavior*, indeed, to our survival. Predictions allow people to plan their future by *adapting themselves* to the predicted sequence of events. Explanations are concerned with finding causes ([sec. 6.2](#)); they provide people with *knowledge of causes*. Causal knowledge gives people more power than merely predictive knowledge: it enables them to causally intervene in the course of events, and thereby *adapt* the world to suit *their own wishes*. The scientific subgoals of prediction and causal control have been pursued by humanity since the Stone Age. Anyone who can *predict* that nourishing plants are growing in a given area can decide to migrate there and thereby ensure a high reproduction rate for their own tribal group. But those who know the *causes* of the growth of nourishing plants, who know how to plant and cultivate them, need no longer be gatherers, but may become farmers, and practice agriculture instead.

In advanced science, the scientific procedures of explanation and prediction operate not on two but on three levels: actual observation sentences at the lowest level, empirical law hypotheses at the middle level, and scientific theories at the upper level. Between the first and the second, and between the second and the third level, the methodological procedures from M3 and M4 take place, as shown in [fig. 2.3-1](#). This threefold division of scientific work is advantageous for many reasons; in all advanced scientific disciplines, a division of labor has arisen between practical experimenters (levels 1 and 2) and theoreticians (levels 2 and 3). To give an example: In the case of the physical theory of the planetary orbits ([sec. 5.1](#)), level 1 corresponds to the actual astronomic observations (e.g. by Tycho Brahe), level 2 corresponds to the empirical laws of the (elliptical) planetary orbits around the Sun (inductively obtained by Johannes Kepler), and level 3 to Isaac Newton's theoretical explanations of these orbits by the Sun's gravitational attraction together with the planets' centrifugal force.

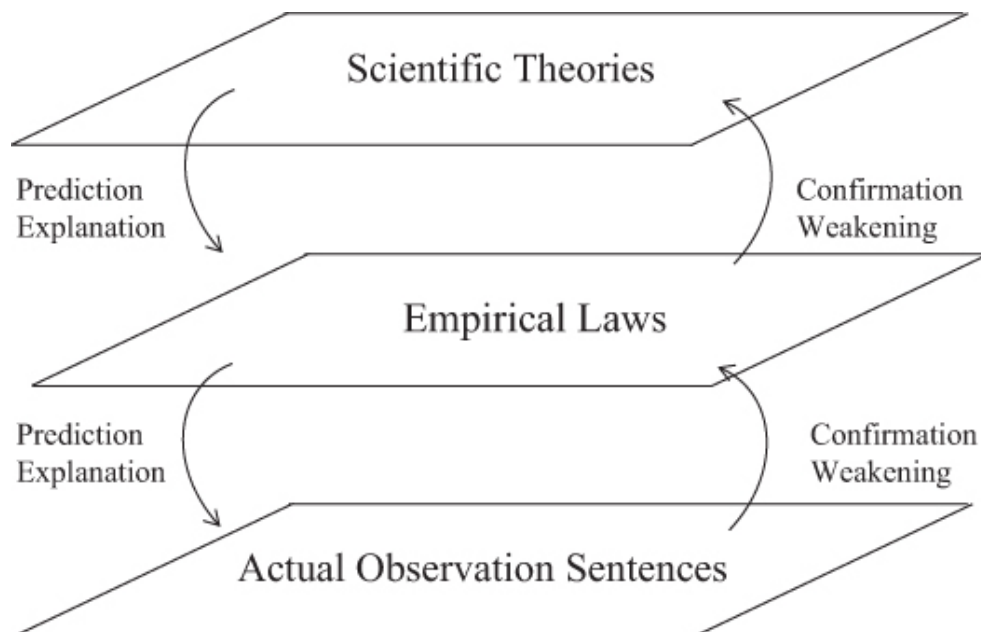


Figure 2.3-1: The three levels of the scientific method. [↗](#)

2.4 Scientific Disciplines and Their Classification

In this chapter, we will examine to what extent our epistemological assumptions E1–5 and the methodological features M1–4 match the existing scientific disciplines. To structure this investigation systematically, we introduce something that is rarely found in textbooks on the philosophy of science: a division of scientific disciplines according to their area of inquiry (see 2.4-1). Our division is certainly imperfect and we attribute to it not a final but merely a heuristic status. Some disciplines are mixed disciplines and are listed under “also;” further special disciplines can be added:

(2.4-1): *Scientific disciplines, classified according to area of inquiry :*

Sciences...

1. ... *of nature* : physics, chemistry, biology, geology, medicine (astronomy, cosmology, geography, paleontology, history of biological evolution)
2. ... *of technology* : mechanical and electrical engineering ..., also: computer science
3. ... *of human beings* : psychology (also: education, medicine, cognitive science)
4. ... *of (human) society* : sociology, economics, political science (also: anthropology, ethnology, geography)
5. ... *of (human) history* : history (also: anthropology, ethnology, philosophy as history of ideas)

6. ... *of human cultural (mental, social) artifacts* : legal sciences, linguistics, literary science, sciences of fine arts and music, art history, media studies (also: education, religious studies)
7. ... *of formal structures (formal sciences)* : mathematics (logic, statistics, theoretical computer science, system theory...), formal methodology and philosophy of science
8. ... *of the general foundations of human ideas* : philosophy (epistemology, philosophy of science and theoretical philosophy; ethics, esthetics, and practical philosophy)
9. ... *of God* : theology (also: religious studies)

Our classification has the advantage of not being tainted by partly value-based categories such as (natural) *science*, *humanities*, or *real science* (etc.). If we try to apply these catch-all categories to our area-of-inquiry-based classification, we recognize that there is no clear consensus about what goes where, and that different systems of classification often have more to do with the politics than the philosophy of science. Group 1, for example, covers the core *natural sciences*. But mathematics too sees itself as a natural science, although it belongs to the category of *formal sciences*; moreover, in the meantime, mathematics is not only applied in natural sciences but also in economics and social sciences. Of late, psychology also sees itself predominantly as a natural science, although it could equally be assigned to the humanities or social sciences.

Groups 5 and 6 comprise the core of the traditional *humanities*. Philosophy, though, is just as close to the natural sciences as it is to the humanities, although in contemporary universities, it is usually placed in the faculty of arts and humanities. In [sec. 1.3.8 – 1.3.9](#), we explained that since the 19th century, the traditional humanities were increasingly understood as a counterpart or even an opposition to the natural sciences. The term *humanities and social sciences* was coined in the 1960s and intended to serve as a modernized counterpart to the natural sciences. Whether this terminological group covers only groups 3 and 4, or also groups 5 and 6, was not terminologically fixed.

In the last two decades, the new terminological group of the *cultural sciences* was formed, in part as a reaction to contemporary tendencies in psychology and social sciences to increasingly use methods of natural sciences. The term “cultural sciences” is predominantly used in continental Europe and is understood in a broader sense than the Anglo-Saxon “cultural studies,” though its extension is not precisely defined. The frequently supposed opposition between natural and cultural sciences is an instance of what Snow ([1960](#)) has called the two ‘cultures’ in the current scientific age. In this book, we intend to refute the view that there exists a principal methodological *gap* between the sciences of nature and culture; this gap, where it exists, is more a matter of different value-attitudes and worldviews than a matter of unbridgeable differences in method.

In the term *factual sciences* (in German “Realwissenschaft”), all sciences are brought together that have a part of the *real* world as an area of inquiry; so, factual sciences include all groups apart from the formal sciences. Theology can only be seen as a factual science if we assume the reality of God, as this reality cannot be scientifically demonstrated, its inclusion here is questionable.

2.4.1 The Special Status of Mathematical Sciences

First, we have to explain the special status of group 7, mathematics, and formal sciences. In their totality, assumptions E1–5 and features M1–4 only make sense for the factual-empirical sciences. Mathematics and formal sciences do not have a domain of inquiry that consists of *concrete* and empirically accessible objects, but rather deal with formal and *abstract* structures. Mathematics, then, has no empirical basis, and therefore assumption E5 and methodological features M2 and M4, which require an empirical basis, do not apply. This does not mean that mathematics is a speculative discipline, which makes statements about reality with no empirical basis: pure mathematics makes no statements at all about an (empirically accessible) *reality*. All it does is develop formal structures and models, characterize them by using abstract axioms and definitions which are as general as possible, and derive logical consequences from them. It would be nonsense, for example, to say that the axiomatic system of real numbers or of Euclidean geometry was per se “true” or “false.” Such claims can only be made when mathematical terms are interpreted empirically, for example, when real locations or time points are assigned to real numbers, or real distances to mathematical vectors. But then geometry is no longer a mathematical theory, but becomes a physical one. Einstein once formulated it like this: “As far as the theorems of mathematics relate to reality, they are not certain; and as far as they are certain, they do not relate to reality” (cf. [Carnap 1967](#), 56). So, we can say: exactly *because* mathematics

ignores the empirical interpretation of its models, it is the strictest of all sciences. To be sure, mathematics concerns itself primarily with those structural models which are applicable to certain areas of reality. Seen in this way, mathematical disciplines can be more or less *fruitful* for the real sciences, but they cannot be true or false in the sense of the factual sciences.

Assumptions E2–4 and the methodological feature M1 can be applied to mathematics without restriction (the assumption of fallibility E2 insofar as even mathematical proofs are not guaranteed to be error-free). Assumption E1 of minimal realism applies only in an indirect way, insofar as mathematical structures have the purpose of being applied to the factual sciences. The methodological component M3 applies insofar as mathematicians are interested in the derivation of the various theorems, which express important properties of their models. Such derivations may be termed mathematical *explanations*, as long as it is clear that these are not causal explanations of real facts (cf. [Weingartner 1978](#), 102; [Pincock 2015](#); [Kuorikoski 2021](#)).

Let us not neglect the fact that the *philosophy of mathematics* includes opposing views on the nature of mathematical objects (cf. [Benacerraf/Putnam 1964](#); [Shapiro 2000](#); [Linnebo 2020](#)). According to the widespread view which is shared by us, mathematical objects are abstract *conceptual* objects, that is, conceptual constructions. Real objects, for example, five red roses, can *exemplify* mathematical objects, for example, the number five, but can never be identical with them. This view, called *conceptual structuralism*, was worked out in detail by the group of mathematicians around Bourbaki ([1961](#)) and philosophers of mathematics like Field ([1980](#)). The opposing standpoint in the philosophy of mathematics is *mathematical realism*, which assumes that mathematical objects exist *per se*. This position has been defended, for example, by Frege ([1918](#), 43) and Gödel ([1947](#), 271f). Also in this position, mathematical entities are seen as abstract and *platonic* objects, not located in the reality of space and time, but in a “platonic world of mathematical ideas” which is accessible not empirically, but only through mathematical “intuition.” An argument in favor of mathematical structuralism is Ockham’s razor (cf. 5.8.7): structuralism explains mathematical phenomena equally well as realism, but without the assumption of strange platonic entities. Another position is mathematical naturalism ([Maddy 1997](#)), according to which mathematical theories resemble scientific theories because they can fit reality more or less well in their applications. (The analogous view in the philosophy of logic has been called “anti-exceptionalism;” cf. [Hjortland 2017](#).) Against this view, it must be objected that in mathematical theories the reference to factual applications is fundamentally left open, whereas in empirical theories this reference is semantically fixed. As a result, empirical theories can be tested through experiments, which is undoubtedly *not* the case for mathematical theories.

In summary, mathematics and the other formal sciences are a *specialization* within the whole system of science, dealing with formal models, algorithms, and methods. This specialization has shown itself to be particularly *useful* in the history of science. Put in a nutshell, the information that the mathematician passes to the empirical scientist is this: “*if* your empirical domain satisfies the axioms of this-and-this mathematical structure, *then* it must also satisfy these-and-these theorems about this structure that we have proved; we grant you the satisfaction of the theorems, provided *you* grant us the satisfaction of the axioms.”

So, we characterize the formal sciences as *prerequisite and auxiliary sciences*, which furnish the factual sciences with formal foundations and structural models with known properties.

2.4.2 Testing the Common Foundations of Science

Having clarified the special status of mathematics, let us now ask: with the *exception* of the formal sciences, to what degree do the assumptions E1–5 and the methodological features M1–4 comply with the listed groups of disciplines? Obviously, E1–5 and M1–4 fit all disciplines in groups 1–3 (natural sciences, technical disciplines, psychology). There is also no reason not to apply them to the social sciences and historical disciplines in groups 4 and 5. Minimal realism, fallibilism, objectivity and intersubjectivity (E1–3), and logical clarity (E5) are accepted standards in all these disciplines, and the same is true for the search for general hypotheses, explanations, and predictions (M1, M3). In all these disciplines, empirical data are available, for example, in the form of the results of psychological experiments, interview data, economic and demographic data, and in the historical sciences in the form of historical texts and source materials, etc. By this, assumption E4 and building blocks M2 and M4 are also fulfilled. Of course, it is significantly more difficult to find general law hypotheses in social science and history, because of the complexity of their area of inquiry – but all that is required for the applicability of the methodological building blocks to a given

science is that the goals therein can reasonably be *sought after*, which does not necessarily mean that these goals are equally easy to attain in all different disciplines.

The same considerations apply to broad areas of the disciplines in group 6. Here as well, the epistemological assumptions E1–3 and (to a lesser extent) E5 are widely accepted standards. Moreover, these disciplines work with empirical data too; so, assumption E4 and the methodological building block M2 apply. For example, linguistic texts comprise the data for linguistics, and works of art and music from various cultures or stylistic periods are the data for literary criticism, art history, and music studies ([Kreuz/MacNealy 1996](#)). Members of the hermeneutic school raise the objection at this point, that “data” in the social sciences and humanities, for example, speech acts or texts, are always of an *interpretative* nature, so that understanding in the humanities is always subject to an unbreakable *hermeneutic circle* (cf. [Gadamer 1975](#) /2005, 235ff). But a number of philosophers of science have worked on the view that the testing of interpretative hypotheses concerning what an author intended to say proceeds in the same “holistic” manner as the testing of theories in the natural sciences (cf. Stegmüller [1979a](#) , 27–86; [Follesdal 1979](#) ; [Rosenberg 2008](#) , 43f). What hermeneutic philosophers rightly point out is that what one calls ‘data’ in the humanities are usually not raw data, but semantically interpreted data in the form of meaningful linguistic or nonlinguistic acts. However, also what physicists call data are usually theoretically interpreted measurement results, for example, elementary particles observed in the cloud chamber. This does *not* mean that raw or uninterpreted data or stimuli do not exist at all: in our physical example, they exist as condensation droplets in the cloud chamber; and in the interpretation of linguistic acts, they exist in the form of sequences of sounds or letters. Although some advocates of hermeneutics doubt the existence of raw linguistic data, their existence can easily be verified: you just have to hear someone speaking in a completely foreign language; what you will then hear are exactly these raw linguistic data.

The school of *methodological dualism* mentioned in [sec. 1.3.8](#) objects, once more, that there are no general law hypotheses in the humanities. If that were true, then the activities of these disciplines would have to be limited to methodological building block M2, that is to say, to the gathering of singular historical facts according to the “ideographic method.” What is true is that in the humanities and the social sciences, there are as good as no *strict* (or exceptionless) laws. But, as will be worked out later ([sec. 4.5](#) , [6.4.2](#)), there are *statistical* and in particular *normic* law hypotheses (“if A, then normally B”), which are indispensable for explanations and hypothesis testing in these disciplines. Therefore, the methodological components M1, M3, and M4 apply also to the historical sciences and humanities.

To avoid misunderstandings, when we say that the methodological components M1–M4 can be and have been successfully applied in the humanities, we don’t imply that *all* representatives of this group of disciplines make use of these methods. Within the humanities and social sciences, there is still the opposition between proponents and opponents of empirical-analytic methods – we mentioned Snow’s (1960) “two cultures” and Ross’s (1996) thesis of the “science wars” ([sec. 1.3.9](#)). Some scientists in the humanities and social sciences reject these methods simply *because* they are empirical and analytic. What this book nevertheless intends to show is that the methodological features M1–M4 are optimal means to pursue the supreme goal (G) of finding content-rich truths even in the human and social sciences.

There is, however, another phenomenon, found mainly in the humanities, that seems to stand in decisive opposition to the scientific method according to E1–5 and M1–4: namely, the phenomenon of *norms and value judgments*. We illustrate this problem at hand of a subdiscipline of legal sciences: jurisprudence, that is, the science of legal judgment. This discipline examines concrete actions and the facts of a case to determine how they fit into the given legal system, and specifically whether they fall under a certain category of offense or crime (going along with a certain punishment). The philosophy of *legal positivism*, which goes back to [Kelsen \(1960/2002\)](#), held the view that jurisprudence is a value-free systematic-deductive procedure, based only on *positive* law, that is, laws which have been passed by legislators. But Kelsen’s thesis is controversial in the philosophy of law. For example, was the removal of private property from a grocer’s shop *petty larceny* or *theft*? (an example from [Savigny et al. 1976](#) , 101f). Or: does an underwear advertisement poster still conform to ‘public decency,’ or does it already violate ‘the dignity of women?’ (an example from [Hilgendorf 2000](#) , 15, 24). In these and other cases, it does not follow from the stipulations of the law and empirical knowledge of the facts, how the deed of the accused should be classified; and the judges and other interpreters of the law now have the task of *interpreting* the underdefined stipulations of that law. In [Savigny et al. \(1976\)](#) , 106, 144ff), it is shown that these interpretations are often based on *value assumptions* (or assumptions of justness), which cannot be found in legal texts, but rather come from an

‘intuitive legal consciousness.’ Now, a legal scientist can certainly show that *if* certain value assumptions are taken for granted, *then* a certain interpretation of the law is adequate. But how can she determine scientifically *which* value assumptions are justified? Can the legal scientist prove that a specified degree of misery of the accused person is sufficient for us to speak of ‘petty larceny’ instead of theft? Can she scientifically decide whether a billboard advertising underwear is still in accord with ‘common decency’ or is already an affront to the ‘dignity of women’? So the question which arises *not only* in jurisprudence, but also in other disciplines relating to the human or social sciences, is this: can fundamental value assumptions be scientifically justified?

Without further ado, the answer proposed by us is *no*. This does not mean that value- and norm-related issues do not occur in science. For one thing, the de facto values of persons can be *empirically* investigated. Moreover, specific value sentences can be logically deduced from *assumed* fundamental value sentences (and further premises). But a fundamental scientific justification of values is impossible. Our view here is concordant with the general consensus of (empirical) science and does not claim to be original; but our justification of this answer should contain some originality. For the value-neutrality requirement has been criticized so sharply by critics of empirical-analytical philosophy of science that it simply cannot nowadays be taken for granted, but needs to be worked through with great care. This task is addressed in [sec. 2.5](#).

2.4.3 Two Further Classifications of Sciences

The above subdivision of scientific disciplines was given according to area of inquiry. To end this section, we will explain two further possible systems of classification. One methodologically significant classification of the factual sciences follows the “invasiveness” of their empirical discovery methods. The classification is summarized in [fig. 2.4-1](#).

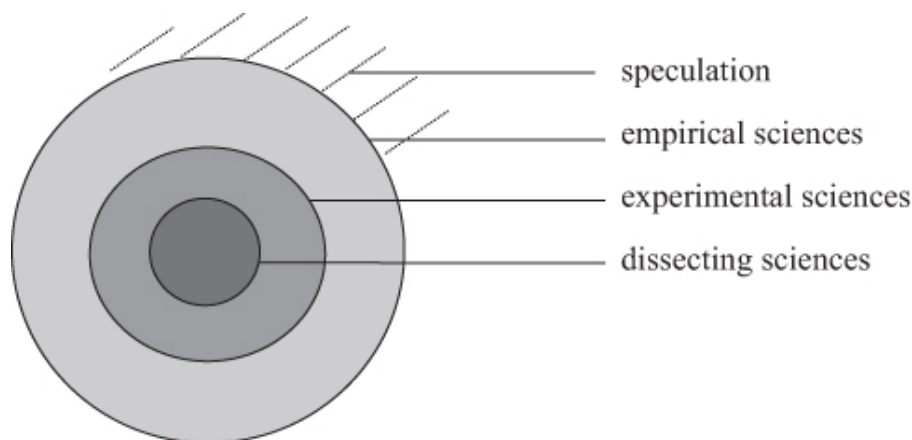


Figure 2.4-1 Classification of the factual sciences. [↗](#)

1. *Empirical sciences* (in the sense of minimal empiricism): By this, we understand all disciplines which refer in some way to empirical data, whether these data were obtained via free field observation, by studying historical records, or in laboratory experiments. To this most extensive group, we assign all disciplines described by assumptions E1–5 and methodological building blocks M1–4, including the humanities. Only those disciplines whose theories are based on pure *speculation* are not in this group – as for example mysticism, esotericism, speculative metaphysics, or the dogmatic and creed-based components of theology. Based on our epistemological assumptions, it is doubtful whether these are factual sciences at all.
2. *Experimental sciences*: This *subgroup* of empirical disciplines obtains its data from *controlled experiments*. The significance of controlled experiments as opposed to free field observation is that in an experiment, certain sets of conditions can be deliberately created, which give us information about the causal relevance or irrelevance of factors (see [sec. 4.2 – 4.4](#)), while field observers might have to wait ‘forever’ for a lucky coincidence to come up with the exact set of conditions required. For this reason, experimental sciences can explore their domain of inquiry much faster and more effectively

than sciences which rely on mere observation or passive reception of historical data. The experimental criterion separates the disciplines in groups 1–3 from the ones in groups 4–6, at least in tendency. For it is hardly possible, for both technical and ethical reasons, to perform experiments with large social groups, institutions, or legal systems. Of course, ethical concerns are already present in biological or psychological experiments, but here they are usually not strong enough to exclude them, only to restrict them.

3. *Dissecting sciences* : A significant subgroup of the experimental sciences are those which I will call – for want of a better word – the “dissecting” sciences. By this, I mean sciences, which not only expose their objects to various experimental conditions in order to analyze them intellectually, but which also *materially analyze* their objects, that is, take them apart or dissect them – perhaps in order to synthesize them afterward, that is, to put them back together again, provided this is still possible. I suggest that it is this dissecting character that constitutes the essential characteristic of *classical* natural sciences like physics, chemistry, biology, and medicine, in contrast say to experimental behavioral research or psychology.

Dissection of objects goes far beyond experimentation at the input-output level insofar as it makes the objects’ inner structure accessible to observation and experimentation. Chemistry and biochemistry have learned to break down items from our environment into their molecular and atomic constituents (analytical chemistry) and to put them back together again (synthetic chemistry). Nuclear physics has done the same at a *subatomic* level, here involving much greater amounts of energy. The analysis and synthesis of living creatures being worked on by science today – from genetic engineering and neuroengineering to artificial intelligence and robotics – could bring us to a point in the future where it is possible to genetically alter or even create a human or android being. Some people think these possibilities are terrifying; others that they are breathtaking. Here, as always, we should remember: scientific discoveries may be used for various purposes, both “good” and “evil.” But certainly, the possible risks and dangers going along with these technological possibilities are getting bigger and bigger. Therefore, the *ethics of science*, whose purpose is to set limits to the technological realization of the scientifically possible, is going to become more significant.

Independently of this, scientific disciplines may also be divided up according to the degree to which their concepts and *methodologies* are quantitative and mathematical, as follows:

Graduated division of scientific methodologies by increasing degree (←) of logical-mathematical and quantitative precision



Quantitative methods are on the left and *qualitative* methods on the right side of this spectrum. Note that on both sides we are looking at *scientific* methods – also qualitative hypotheses and methods can be pursued in a logically clean and empirically controlled way (cf. [sec. 3.1.4](#) ; [Denzin/Lincoln 2000](#)). The ideological polarization between quantitative and qualitative methods that is held by some qualitative researchers (e.g. [Lamnek 1988](#)) appears unnecessary and exaggerated. The strength of qualitative methods (e.g. case studies, narrative interviews) lies in advance of quantitative methods – in the exploration of relevant parameters and the generation of promising hypotheses. But a qualitative exploration has to be followed up by a quantitative-statistical analysis, as this is the only reliable way to test the generality of one’s hypothesis. That qualitative and quantitative methods are *complementary* in this sense is a widely held view among empirical researchers in the social sciences (cf. [Lazarsfeld/Barton 1951](#) ; [Giddens 1982](#) , 328); however, this view is not uncontroversial (cf. [Flyvbjerg 2006](#)).

2.5 The Question of Value Neutrality

2.5.1 Max Weber's Value-Freedom Postulate

The first detailed work on the value-neutrality requirement goes back to one of the main founders of sociology, namely, *Max Weber* (1864–1920). Weber was involved in a controversy with Gustav Schmoller, who propagated the view that social science should work out moral value judgments relating to a desired social order. Weber spoke out strongly against the pseudo-objectivity of such “podium values,” that is, value judgments presented with ‘scientific’ authority: Objectivity, said Weber, could only be reached in science if scientists limit themselves to factual, descriptive statements, and separate these clearly from their own values ([Weber 1922](#) /1949, ch. II and X).

The main reason for the self-imposed limitations of science to descriptive judgments, recognized by Weber, may be summarized thus: *Values are not properties inherent to objects themselves, but rather are based on subjective interpretations by us humans. Norms are also not objective facts, but man-made requirements, designed to aid the realization of certain values.* In the case of testing value and norm sentences, there is nothing which corresponds to an *empirical foundation* according to assumption E4 and methodological building blocks M2 and M4. At the same time, value and norm sentences are by no means purely formal claims, which might be justifiable by means of pure logic. There is therefore no objectivity in the sense of empirical science when it comes to norms and values. Ultimately, the *decision* for or against certain values or norms is a question of personal *freedom*.

This explanation of the value-freedom postulate is *prima facie* not only theoretically, but also practically plausible. Everyone knows from experience that finding a consensus on values between members of different cultures is much more difficult than finding a consensus on factual issues. Many strict Islamists would not take offense at Western science; the Western value system, however, seems twisted in their eyes. Nonetheless, the value-freedom requirement was massively criticized from the 1960s onward, especially by neo-Marxist movements such as the *critical theory*, and in the German-speaking area the Weber-Schmoller controversy flared up once more as the “positivism dispute” (cf. [Frisby 1972](#) ; [Adorno et al. 1976](#) ; [Dahms 1994](#)) and the “value judgment dispute” ([Albert/Topitsch 1971](#)).

Proponents of critical theory raised the objection that a purely descriptive or “positivist” method in social science had the effect of simply theoretically reproducing the status quo of oppressive social structures and thereby implicitly sanctioning them, while *critical science* must concern itself with *changing* oppressive social conditions ([Habermas 1966](#)).

This accusation is based on a fatal *misunderstanding* : descriptive research, of course, also includes research into natural laws, and thereby research into that which is naturally, technically, or practically *possible* and *impossible* . So, in fact, empirical social science is the very thing which can show us how social conditions may be *changed* , thereby contributing to the emancipation demanded by representatives of critical theory. In fact, almost all the “neopositivists” of the Vienna Circle had a *de facto* progressive political orientation and saw much of their work as being connected with social progress ([Dahms 1994](#) , 125; [Stadler 2001](#)). The only difference is that critical-empirical social science, as defended by Popper and Albert against Horkheimer and Adorno in the “positivism debate,” restricts itself to pointing out the *possibilities* for social change, and leaves normative decisions to the process of *democratic opinion-formation* , rather than wanting to justify these in a quasi-scientific way.

Even if the main accusation of critical theory was off the mark, the period that followed saw a number of legitimate objections being raised, which are in need of a more subtle discussion. Not only critical theorists (cf. [Dahms 1994](#) , 375), but also younger analytic philosophers warned that questions of values and norms should on no account be handed over to *irrationality* . Indeed, a rational treatment of value questions is both possible and necessary in society, and science has an important role to play in this regard. In this regard too, it is advisable to turn to Max Weber, who made subtle refinements to the value-neutrality requirement, in the face of just such objections:

First , according to Weber, it is absolutely possible for a social scientist to do research on the factual presence of value and norm systems, *without* passing judgment herself ([Weber 1922](#) /1949, 11f). To study polygamy in Islam, it is neither necessary to be an advocate nor to be an opponent of it. The statement that a

person or a society *believes* in certain values is not in itself a value statement, but a *descriptive* claim, whose confirmation requires only empirical data.

Second, also according to Weber ([1922](#) /1949, 21f), the scientist can identify *logical relationships* among value or norm sentences. Relevant to this task are the – meanwhile highly developed – disciplines of analytic ethics (e.g. [Frankena 1963](#)), deontic logic (e.g. [Aqvist 1984](#)), and decision theory (e.g. [Raiffa 1968](#)).

Third and most importantly, with the help of *means-end inferences*, the scientist can infer from *given* norms and descriptive knowledge a series of derived norms about optimal means of reaching accepted ends and pass these on as recommendations to society and politics. Because this represents the most important function of science for society, the next section is devoted to it.

2.5.2 Value Neutrality and Means-End Inferences

Finding suitable means for given ends is an important task of descriptive science. It obeys the following logical form:

(Def. 2.5-1) *The general scheme of means-end inferences* :

Descriptive means-end premise : M is in the given circumstances C a necessary – or alternatively an optimal* – means for the realization of end E.

Thus – hypothetical value- resp. norm conclusion : If end E is accepted as a fundamental value (or norm), then the realization of means M is a derived value (or norm).

Means-end inferences are accepted by most ethical theories as analytically *valid* (the asterisk* is explained below). The conclusion can be formulated for value statements (“P is or is not valuable”) and normative statements (“P ought or ought not to be done”). Both are analytically related, which can roughly be expressed by the formula: “The good ought to be done and the bad ought to be avoided” (for details, see [Schurz 1997](#) sec. 11.2–11.3). Value judgments cannot be analytically reduced to factual statements about the value sentiments of some or many people (as [Anderson 2004](#) proposed), but rather possess *normative force*.

The recommendation of means in the conclusion is not understood in a categorical, but in a hypothetical or conditional sense. That is, it is relativized to the assumed ‘fundamental’ or initial values (ends) E: *If* such-and-such fundamental values are assumed, *then* certain action recommendations follow as derived instrumental values. For conceptual clarification: A *categorical* value statement has the form of a basic value sentence “P is (or is not) valuable;” a *hypothetical* or conditional value statement takes the form of an if-then sentence “If P is (not) valuable, then Q is (not) valuable” (this distinction goes back to [Kant 1785](#)). A categorical value statement is *fundamental* if it cannot be derived from other categorical value statements using means-end inferences; otherwise, it is *derived*.

It is crucial for the value neutrality of science that recommendations of means are formulated hypothetically rather than categorically. Only in this way can a recommendation of means follow from factual knowledge and thus remain value-neutral itself, and only in this way can objectively value-neutral science be reconciled with societal value engagement. Value judgments derived hypothetically by science are also referred to as *instrumental* rationality; however, we avoid this term, as it implies a certain devaluation (“merely instrumental”), whereas what we mean is a genuine value engagement – only formulated hypothetically in order to clearly *separate* the factual component from the value component of the statement.

Another advantage is that, through the use of rich scientific knowledge, a large number of derived norms can be obtained from just a few fundamental norms. For example, the EU climate goal of becoming carbon-neutral by 2050 entails a multitude of recommendations of means. In this way, the burden of justification can be shifted – if not entirely, then to a large extent – from the normative to the descriptive.

In a means-end inference, it is assumed that M is either a necessary means – that is, every realization of the goal requires the realization of M – or an optimal means, meaning that among all practically possible alternatives, M maximizes total utility with regard to goal realization and side effects. Strictly speaking, the definition of an optimal means requires a listing and cost-benefit evaluation of all alternative means. In order for the hypothetical recommendation of optimal means to remain value-neutral, this cost-benefit evaluation must be included as a fundamental value assumption in the antecedent (the “if”-part) of the recommendation

– which is indicated by the asterisk. The precise reconstruction is therefore more complex than presented in [def. 2.5-1](#) (see [Schurz 2024b](#), ch. 3).

If M, on the other hand, is a merely *sufficient* means for an end E, then the means-end inference is not valid, because many sufficient means cause harm that outweighs the benefit of achieving the end. For example, to achieve the end of letting fresh air into a room, smashing a hole through the wall is a sufficient but not recommendable means; by contrast, opening a window is an optimal means, and a necessary means is some sort of opening to the outdoors.

Neither the inference to necessary nor to optimal means implies the mistaken thesis that “the end justifies the means:” if the costs of every possible realization of the end E exceed its benefits, then E should not be realized and can no longer be considered an acceptable fundamental value (given a consequentialist understanding of value).

The scientists or experts should take the fundamental norms from the knowledge *users* – that is, the public or their political representatives. Based on these norms, the scientists can then derive recommendations of means from scientific knowledge and pass them back to the knowledge users. Let us illustrate this by means of educational *teaching theories* (cf. [Flanders 1970](#) ; [Patry 1991](#)). Let us take it that *cognitive success* (i.e. mastery of material to be learned) is the fundamental goal of teaching. Based on empirical research, an education scientist may reach the descriptive hypothesis that cognitive success may be maximized if between 70% and 80% of teaching time is teacher-guided. The education scientist passes on this conclusion to teachers as a *derived* norm. This procedure is, however, only in accordance with the value-neutrality requirement if the means recommendation is formulated *hypothetically* : if your goal is to maximize cognitive success, then 70%–80% of class time should be teacher- guided.

If the explicit reference to assumed values is neglected, ideologically worrying consequences can result. Assume that in our example, the teachers have different fundamental goals than the scientists; they see *social learning* as an equally important goal to cognitive success. Let it be empirically shown, perhaps even by the same scientist, that the optimal means to reach this purpose is a maximum of just 60% of teacher-guided class time, alongside 40% of group activity. What if the scientist in our example, however, passes on her recommendation categorically rather than hypothetically, for example, she asserts *80% of class time should be teacher-guided* , and teachers believe him? Now the scientist is *manipulating* the teachers, because she is concealing the fact that her means recommendation is valid only for *her own* value assumptions. The teachers, who *do not* share these values, are being badly advised on how to reach their own chosen ends.

The same is true of many other situations, from medicine (a doctor advises a patient) to law (a lawyer advises a married couple about to get divorced). It is highly important in all such cases that the scientific “expert” makes the value assumptions *explicit* that she assumes in her means recommendations, as this is the only way that the “customer” being advised can find out whether these correspond to his own values. With these examples, we *emphatically* point out that the value-neutrality requirement does not have a dogmatic, but rather a *democratic* and *emancipatory* function. In [sec. 2.10](#) , we will defend this thesis in the context of contemporary debates on values and their role in science.

In summary, the central conclusion of this section is that value neutrality and value engagement in science are indeed compatible. The oversimplification of the value-neutrality requirement into the “value-free doctrine,” according to which values have “no place” in the empirical sciences, is therefore a serious misrepresentation that often serves as an easily criticized straw man for opponents of value neutrality. Sciences can and should use means-end reasoning to derive values from externally given fundamental values and pass them on to knowledge users in the form of hypothetical recommendations of means. Only the *justification of fundamental values* lies outside the scope of objective science.

2.5.3 Explication of the Value-Neutrality Requirement

To reach the final formulation of our value-neutrality requirement, we need to take two further steps of clarification.

1. Many authors have argued that science cannot be value free in an absolute sense, because the whole enterprise of “science” itself rests on certain goals and values: the so-called *science-internal* or epistemic values (e.g. [Schmidt 1959](#) , 648ff; [Doppelt 2007](#) , 190ff; [Sober 2007](#) , 110). In [sec. 2.1](#) , we identified the search for general and content-rich truths as the supreme epistemic goal of all sciences.

This fundamental internal value of science is accepted by all scientific disciplines, from which a range of derived epistemic values follow as means to achieve this supreme epistemic value (see [Doppelt 2007](#)) – in particular, the four methodological characteristics mentioned in [sec. 2.3](#). The value-neutrality thesis must therefore be restricted to fundamental *science-external* or non-epistemic value assumptions.

2. Is the restriction to fundamental science-external values and norms enough to give us a sustainable version of the value-neutrality requirement? *No* – insofar as one views science not merely as a process of justification, but as a *social process*. The process of scientific discovery is usually divided into three phases: the *context of origination* (CO) (also called “context of discovery”), in which the scientifically relevant problems are first identified and defined, the *context of justification* (CJ), in which data are gathered and systems of hypotheses generated and tested, and the *context of application* (CA), in which well-tested discoveries are used for various different social purposes. In both the CO and the CA, non-epistemic values inevitably play a role. Decisions about which research questions are considered important enough to pursue are also influenced by non-epistemic interests – such as those of funding bodies or the research team. Even more strongly, non-epistemic interests determine the purposes for which scientific knowledge is used in the CA. What the value-neutrality thesis demands relates exclusively to the CJ: in CJ, science-external values must not play any role. In conclusion, we can formulate and schematically represent the value-neutrality requirement as follows (see [fig. 2.5-1](#)):

(Prop. 2.5-1) *Value-neutrality requirement (V)* : A specific realm of scientific activity, namely its context of justification, should be free from fundamental science-external value assumptions.

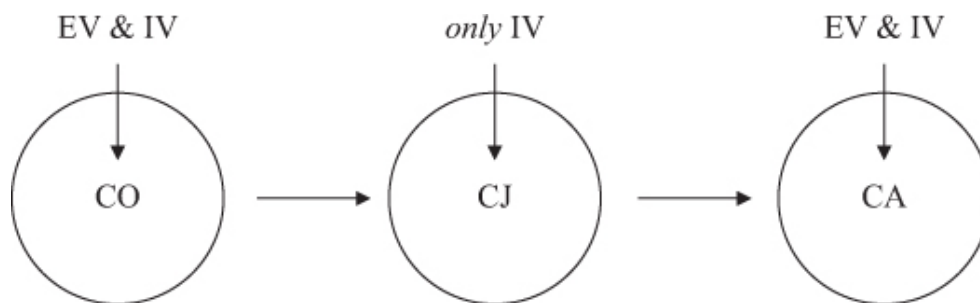


Figure 2.5-1 Schematic representation of the value-neutrality requirement. [↗](#)

Legend: CO – context of origination, CJ – context of justification, CA – context of application; IV – science-internal values, EV – science-external values.

The value-neutrality requirement will hereafter be abbreviated as (V). Note that (V) is a normative requirement, not a factual claim about the actual judgment practices of all or most scientists. There is broad acceptance of the value-neutrality requirement in the sciences, but as noted in [sec. 1.2.9](#) and [2.4.2](#), the requirement is rejected by a significant number of scholars in the humanities and cultural studies (keyword “science wars”). The content of the value-neutrality requirement can also be expressed as a thesis without an explicit normative formulation as follows: the *justificatory status* of scientific claims is *independent* of non-epistemic values. However, even in this formulation, it remains a statement about the epistemic-normative justification status of scientific judgments – and not about the actual judgment practices of scientists.

In our modern, value-pluralistic information societies, politicians and the media rely on scientific legitimations. As a result, scientists face a constant pressure, or at least the temptation, to present recommendations of means in the form of categorical value judgments. Scientists should resist this temptation, as doing so ultimately undermines public trust in science, as recent empirical studies have shown (see [Angeli et al. 2021](#); [Intemann 2024](#), 10). The distinction between factual and value judgments is an essential prerequisite for any credible differentiation between true facts and fake facts, and thus for countering science denial and conspiracy thinking in our time. Only objectivity and value neutrality can enable stable scientific progress through the politically and culturally shifting phases of history.

2.6 The Demarcation Problem and the Methodological Unity of Sciences

2.6.1 Demarcation

If we add the value-neutrality requirement (V) to the epistemic goal (G), the epistemological assumptions (E1–5) and methodological features (M1–4), we can thereby characterize the boundaries of the scientific. As long as legal science, for example, analyzes empirically occurring legal systems, or establishes logical and means-end relationships between norms, it remains a descriptive science. The logical analysis of ethical statements carried out by analytic ethics and meta-ethics also fits easily into the logical-descriptive scientific program. Literary and art theory, which deal with the aesthetic effects of various literary and artistic genres, fit into the framework of empirical-descriptive science, as long as aesthetic effect refers to a *factual* effect on people. Only absolute statements about aesthetic values, which are not relativized to a factually given artistic style, go beyond the scope of descriptive science. This criticism also applies to the tendencies within humanistic psychology and pedagogy which place values like personal development and dignity above objective analysis ([Bühler/Allen 1973](#)). Religious studies, as long as they reconstruct factually occurring religious systems, belong to the descriptive scientific program, while theologies that dogmatically assume the truth of a certain religious, “creed” are departing doubly from the realm of science: first, by assuming the truth of speculations about the nature and existence of God and second by making fundamental value assumptions.

In summary, we can demarcate the boundaries of science as follows: *Sciences in the sense used here are all empirical disciplines in the broad sense (according to G, E1–5, M1–4 and V), together with their formal auxiliary sciences and their means-end-based applied disciplines.*

This characterization of “science” gives us a demarcation between scientific and nonscientific disciplines and fields of mental activity. So, it serves as a suggested solution of the *demarcation problem* (cf. [sec. 1.1](#)). We do not understand this demarcation problem, though, in the same way as logical positivism or critical rationalism did. *First* , we do not identify the nonscientific with the cognitively nonsensical (as e.g. [Carnap 1928b](#)). This identification is unacceptable, because empirically vacuous sentences like “God exists” or fundamental norm sentences like “Help your neighbor” may of course be *understood* , and so have a communicable *meaning* . *Second* , we do not devalue such sentences; we classify them as nonscientific, simply because there is no way to test them empirically. *Third* , we see the significance of the demarcation problem differently from Carnap. Our proposed solution of the demarcation problem does not pursue the inner-philosophical concern of eliminating a part of philosophy called “metaphysics.” Rather, it serves (A) a societal demarcation aim and (B) an evolutionary explanatory concern:

Point A concerns the public debate on the question which kinds of information are sufficiently objective and confirmed to be taught in public schools, and which are not. This requires demarcation criteria between science and religious ideology. As an example, we mentioned in [sec. 1.1](#) two famous cases in which judges based their decisions against the teaching of creationism in public schools on demarcation criteria developed by philosophers of science. However, there is a controversy among philosophers of science themselves as to what the right demarcation criteria should be. Ladyman and Ross ([2007](#) , 30) prefer sociological over logical demarcation criteria. But sociological demarcation criteria are insufficient. For example, Pennock ([2011](#) , 193ff) reports that Judge Jones suggested the application of peer-review methods as a social demarcation criterion. The reaction of Behe, a prominent intelligent design defender, was to promote peer reviews within the creationist camp. Behe pointed to a couple of creationist writers that indeed had been peer-reviewed – not by naturalistic scientists, but by other intellectual creationists. But the consent which a creationist writing earns from other creationists is clearly not an indication of its scientific character.

Point B concerns the *explanation* of the *evolutionary impact* of the sciences. Looking back to the last one or two millennia of human development, science has made extremely rapid progress, compared to ethics, religion, and arts which have developed slowly, at least in those domains where they have not been driven by external technological change. The factual differences in the level of evolutionary impact of the sciences in comparison to ethics, religion or art, and the degree in which they have changed our life, are *so* marked, that they positively beg to be explained. Anyone who wants to understand the powerful independent dynamics of

science and technology has to find out, *what it is* that separates science so strongly from nonscientific systems of ideas. In conclusion: Any philosophy of science which wants to be of practical significance must face up to the demarcation problem.

2.6.2 Methodological Unity instead of Physical Reductionism

Our demarcation account also contributes to the question of the *unity of science*. However, our understanding of the unity of sciences is not ontological, but methodological in nature, and therefore different from traditional accounts. These accounts were predominantly based on *reduction* attempts, that is, on attempts to provide an explanatory or ontological reduction from 'higher-level' or 'special' sciences to *physics*. There are, however, weaker and stronger sorts of reduction accounts. The career of 20th-century reduction accounts started with the program of *physicalism* of Otto Neurath, a prominent logical empiricist, who announced this program as a means to establish the unity of sciences. Neurath's physicalism was weaker than it sounds, insofar, it consisted only in the requirement that all scientific sentences should be translatable into statements about concrete spatiotemporal things and properties ([Neurath 1934](#) ; [Dahms 1994](#) , 170).

A much stronger form of reductionism had been developed in the 1950s and 1960s: the program of *physical micro-reduction*, whose goal was an ontological or explanatory reduction of the laws of special sciences to laws about underlying physical microstructure. ¹

There exist different concepts of reduction. In the standard account of Hempel ([1969](#)) and Nagel ([1970](#) , 125), a physical reduction of a law describing the behavior of a complex system consists in the derivation of this law, or an approximative version of it, from fundamental laws of physics together with contingent boundary conditions, and suitable correspondence principles that translate higher-level concepts into physical concepts. More recent notions of reduction are discussed in van Riel and van Gulick (2025); as these discussions suggest, the core ideas of the Hempel-Nagel concept of reduction can still be considered well-established.

In spite of some positive examples (e.g. the micro-reduction of temperature to mean kinetic energy), this reduction program was not especially successful. It does not describe the way in which special sciences typically proceed. The objects of the special sciences are more-or-less complex systems. At every level of (at least moderately high) complexity, special sciences have discovered their own laws, which are called *system laws* in [sec. 6.5.1](#). Examples are the chemical laws describing the behavior of molecules, biological laws about the functioning of organisms, psychological laws describing human behavior, etc. There exist three arguments, which show that the laws of special sciences are typically *not* reducible to physics, at least not fully (cf. [Cartwright 1983](#) , 104f, 113f):

1. The argument from *hypercomplexity* : The fundamental physical law governing a specific physical system is expressed by its Schroedinger equation. However, already for moderately complex microsystems such as a medium-sized molecule (as opposed to a simple atom), their Schroedinger equation is far too complex to be solvable without radical simplification assumptions. These simplification assumptions do *not* come from fundamental physics but from empirical chemical knowledge. For these reasons, most chemists are convinced that chemistry will never be *fully* reducible to physics (cf. [Hendry 2008](#) , 525ff.; [Mortimer 1986](#)). Much more is this true for the laws that describe highly complex systems such as biological systems: here already the formulation of an approximately complete Schroedinger equation is impossible and attempts to derive biological normal-case laws ([sec. 4.5.2](#)) from fundamental physics and factual descriptions of the physical microstructure of an organism are hopeless. Rather, biological laws are typically discovered by empirical-inductive means. This does not exclude the existence of *partial* reduction relations, for example, between biochemistry and biology or neuroscience.
2. The argument from *multiple realizations* : Empirically, identical macro- properties of a complex system, such as elasticity or intelligence, can be realized by very different physical microsystems ([Fodor 1974](#)). It is hard to find general principles that translate a given macro-property into an exhaustive disjunction of physical micro-properties.
3. The argument from the *systemic properties* : The parts of a complex system are systemic, which means that the properties of these parts in isolation or in a simpler system are different from the properties of the parts within the complex system. An example are the properties of the hydrogen atom H in water

molecules H_2O (O for oxygen), which are different from the properties of the H-atom in isolation, or in the H_2 -molecule. ²

The argument from systemic properties only shows that a so-called horizontal reduction is usually impossible, that is, the explanation of a complex system by the properties of its parts in isolation or in simpler systems. By contrast, “vertical reduction” means that the higher-level properties of a system are reducible to the lower-level properties of its parts *within* the system. Boogerd et al. ([2005](#)) argue in application to biochemical networks that in spite of horizontal irreducibility, vertical reducibility is often possible, and is sufficient for explanations based on biochemical mechanisms. However, since the systemic properties of the parts are determined by the entire structure of the complex system, a micro-reduction of the systemic properties of the parts would lead into an explanatory circle and is therefore not feasible.

These arguments show that the hopes for a physical micro-reduction of the laws of complex systems are dim. For this reason, many philosophers of science have replaced the idea of explanatory reduction by the weaker requirement of *supervenience* ([Hüttemann 2004](#)). Supervenience requires not an explanatory reduction, but merely a “vertical” ontological determination: the physical microstructure of a system *determines* its higher-level properties (at the same given time point). As this determination relation may be infinitely complex, it does not imply the possibility of a reduction by human cognitive resources.

The opposite of supervenience is ontological *dualism*, according to which two physically identical systems may nevertheless differ in some higher-level non-physical properties, for example, “having a mind” according to mind-body dualism. In fact, the supervenience position has been developed by Davidson ([1970](#)) as a naturalistic counter-position to Cartesian mind-body dualism. The supervenience position is fully compatible with the possibility, for which we have argued in this section, that an explanatory reduction of the sciences of complex systems to physics will not be possible in the foreseeable future.

In conclusion, the prospects to establish the unity of sciences by methods of a reduction to fundamental physics are not promising. What one finds along this route is not unity, but *ontological variety*. We suggest that the unity of the sciences should rather be sought by highlighting their epistemic-methodological common ground. The unity of the sciences, if it exists at all, does not exist at the ontological level, but in the form of their *methodological unity*. Of course, there are also important methodological differences between certain disciplines, which will be worked out in several sections of this book ([sec. 1.3](#), 2.4, 3.1, 4.3, 4.5, 5.1, 5.3, 5.5, 6.4, 6.6). But our considerations so far have shown that there exist essential methodological features which are common to all empirical disciplines, despite their differences, and which delineate scientific from nonscientific intellectual activities: the ultimate epistemic goal G, the epistemological assumptions E1–5, the methodological building blocks M1–4, and the value-neutrality requirement V. In this methodological version, the idea of the unity of science can live on in a humbler, but more sustainable version.

2.7 Scientific Inference and Argumentation

In the methodological building blocks M3 and M4, scientific processes of argumentation and inference take place. Thereby, we are dealing with very different processes of inference, which are explained in this section.

2.7.1 Deduction and Induction

When, in accordance with M3, we use a strict law hypothesis together with observation sentences to infer further observation sentences, then we are dealing with a logical-deductive inference. But when, in accordance with M4, we infer from the fact that a large number of actual observation sentences are instances of a law hypothesis, that the law hypothesis is likely to be true, then we have made an inductive inference, or more accurately an inductive generalization. Here are two examples:

(2.7-1)

Deductive inference (logic i.n.s.):

All fish breathe using gills.

Inductive generalization inference:

All fish observed so far (no. 1, 2, ... n)

This animal is a fish.	have breathed using gills.
Therefore, this animal breathes using gills.	Therefore (probably), all fish breathe using gills.
<i>Certain</i> : truth transfer in all possible worlds.	<i>Uncertain</i> : truth transfer only in sufficiently “uniform” possible worlds.

The decisive difference lies in the fact that only deductive inferences transfer the truth from the premises to the conclusion with *certainty*, that is, in all possible “worlds:” *if* the premises (the sentences above the dividing line) are true, then the conclusion (the sentence beneath the dividing line) is also certainly true. Deductive inferences constitute the object of *logic in the narrow sense*. (Some authors place an inductive logic in opposition to deductive logic, for example, [Carnap 1950b](#), sec.43ff; this terminology is disputed, however, see [sec. 5.9.5](#).)

By contrast, inductive generalizations are fundamentally uncertain – the double line in ([2.7-1](#)) indicates this uncertainty. This is because the premises of an inductive generalization tell us only about the cases observed *so far*, while the conclusion generalizes to *all* and in particular to all *future* cases. For this reason, it is also said that inductive inferences are *content expanding*. Only in circumstances or worlds which are sufficiently *uniform*, whose future and past are sufficiently similar, can we reliably infer the truth of an inductive conclusion from the truth of the premises. Nothing can logically guarantee that the future course of events will be sufficiently similar to the course of events observed so far. An extinct fish without gills could, for example, be discovered tomorrow, or a new species of fish without gills could start to evolve through mutation.

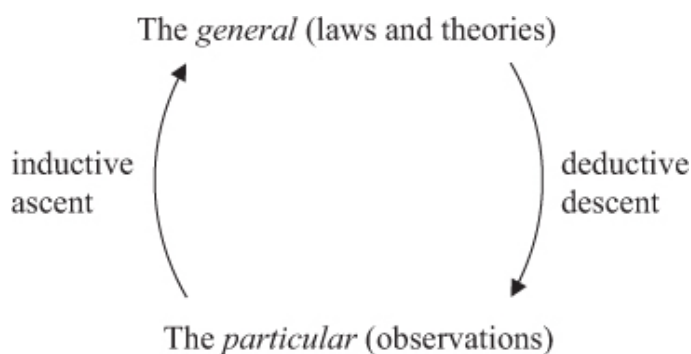


Figure 2.7-1 The inductive-deductive scheme. ↵

This was (very roughly) the main argument of *David Hume*, who was the first philosopher to show that a justification of inductive inference is impossible by means of deductive logic. Hume went on, however, to cast doubt on the rational justifiability of induction as a whole. Hume’s “problem of induction” belongs to the most difficult philosophical problems; a closer look at its subtleties will be given in [sec. 2.11](#).

We have so far explained induction and deduction in the sense of the so-called *inductive-deductive scheme*. According to this scheme, which was already outlined by Aristotle, induction is a method of reasoning (or ascending) from the *particular* to the *general*, and deduction is a method of inferring (or descending) from the general back to the particular (see [fig. 2.7-1](#)).

Although this scheme includes some important methodological features of the sciences, it is *oversimplified* or even incorrect in several respects. For one thing, the binary division of types of inferences on which it rests is *too narrow* – there are far more types of inference, which we will briefly list here. Where the general is concerned, we must distinguish between *strict* (or deterministic) and *statistical* generalizations.

1. Types of deductive inferences :

1.1 From the (strictly) general to the particular – for example, All Fs are Gs, this is an F, therefore: this is a G.

1.2 From the general to the general – for example, All Fs are Gs, All Gs are Hs, therefore: all Fs are Hs.

1.3 *From the particular to the particular* – for example, This is an F or a G, this is not an F, therefore: this is a G.

By contrast, there are no deductive inferences from the particular to the (essentially) general.

2. *Types of inductive inferences* (see also [Carnap 1950b](#), 207f):

2.1 *From the particular to the general* :

2.1.1 *The strict inductive generalization* : All Fs observed so far were G, therefore (probably), all Fs are Gs.

2.1.2 *The inductive-statistical generalization* : r % of all Fs observed so far were G, therefore (probably), approximately r % of all Fs are Gs.

2.2 *From the particular to the particular – the inductive prediction* : All Fs observed so far were G, therefore, the next F will probably be a G, too.

2.3 *From the general to the particular – the inductive-statistical specialization* : r % of all Fs are Gs, this individual is an F, therefore, we can be r % sure that this individual is a G.

Inductive specializations are also called “direct inference” ([Levi 1977](#)) or “statistical principal principle” ([sec. 3.13.4](#)). Two facts about them are remarkable: First, the strict variant of inductive specializations is deductively valid. Second, like inductive generalizations, inductive-statistical specializations assume uniformity: they are only reliable if our actual observation sample is representative for the total domain.

2.7.2 *Popper and the Significance of Inductive Inference in the Sciences*

The founder of critical rationalism, Karl Popper, advocated the thesis that the induction problem is, on the one hand, unsolvable, but, on the other hand, that empirical science can proceed without inductive inferences altogether (1935/2002, sec. I). In the following, we want to demonstrate, based on a differentiation between *three kinds of induction*, that Popper’s arguments bring to light important insights, but that his radical claim that empirical science could go without induction is not tenable (cf. [Schurz 2002a](#)).

1 *Methodological induction*. This view of induction connects directly with the inductive-deductive scheme explained in [sec. 2.7.1](#) . According to this view, induction is a method of inferring general laws and theories by inductive generalization from particular observations. Popper criticizes the view contained in methodological induction that induction is the central method of discovering or generating law hypotheses and theories. The belief that such a discovery method is at all necessary in science rests, says Popper, on a confusion of the *context of discovery* and the *context of justification* (1935/2002, sec. I.1–2). How scientific hypotheses are generated, be it through induction, intuition, or trial and error, is completely *irrelevant* to the context of justification. Seen like this, methodological induction is *dispensable* in the context of the justification of science.

By contrast, call to mind [sec. 2.1](#) : not all but *some* methods of generating hypotheses may give us a certain measure of justification as well. This is the case, for example, for the inductive generalization inferences explained in [sec. 2.7.1](#) : these inductive inferences are methods of discovering empirical law hypotheses, which at the same time inductively justify these law hypotheses, that is, preliminarily confirm them. More refined inductive discovery algorithms can be found in Langley et al. ([1987](#)) (see also [sec. 4.2.2](#) , 4.3.2). Hypothesis generating procedures which at the same time give us a justification are of course cognitively much more *efficient* than blind trial and error procedures.

In two important points, however, Popper’s criticism of methodological induction is right. (i) Hypotheses do not *have* to be produced by methodological induction – all methods of producing hypotheses are allowed and deserve critical assessment. (ii) Inductive generalization methods work for empirical law hypotheses, but they are no longer sufficient for the construction of scientific theories. Even very simple theories – like, for example, the Galilean law of free fall, which posits the idealization of a friction-free fall – cannot be created by inductive generalization procedures alone, since they contain concepts that transcend the observable – so-called *theoretical concepts*, such as “friction-free movement,” “mass,” or “force.” So Popper’s criticism of

the possibility to create scientific theories by methodological induction (see 1983, 118) is to a large extent legitimate.

2 Logical induction . This view of induction goes back to Carnap ([1950b](#)) and understands induction not as a method of discovery, but as one of justification of scientific hypotheses, that is, as a method of determining the *degree of confirmation* conferred on theories by given observational data. This should happen by the establishment of a system of logical probability, in which the inductive probability of a theory given a body of observational evidence may be determined in a rational way (see [sec. 4.7](#)). Instead of the inductive degree of confirmation, one can alternatively regard the degree of (epistemic) *truthlikeness* of a theory relative to a body of evidence as the central measure of evaluating theories (see [sec. 5.11.1](#)).

Popper's criticism of the view of logical induction now says the following (1983, 19f, 23): while we are able to determine that relative to given observations, a theory T_1 is *better* confirmed or corroborated than another theory T_2 , or that T_1 is closer to the truth than T_2 , we still cannot give an *absolute* value for the probability or truthlikeness of T_1 or T_2 , as tomorrow a new superior theory T_3 may be developed, which completely overshadows both T_1 and T_2 . Popper's argument represents a substantial insight into the *limitations* of logical induction. In order to assign a nonarbitrary *numeric value* to the inductive probability or truthlikeness of a theory T given some evidence, one must have at one's disposal, among other things, a comprehensive list of all relevant *alternative theories* $\{T_1, T_2, \dots\}$. However, to establish such a list in a nonarbitrary manner is, at least in the case of theories, impossible. For any given theory T , there are always indefinitely many alternative theories, as these may contain any number of new theoretical concepts. A measurable linguistic space containing all possible alternative theories simply does not exist (see [Howson 2000](#) , 46). Many theories that were later accepted were still unimaginable a century earlier in the history of science (see [Stanford 2006](#)).

Again, in the case of *empirical* hypotheses, the situation is different. Here, it is possible to establish a finite list of all possible alternative hypotheses in a *limited* observational vocabulary. Let us consider, for example, the set of all interval hypotheses of the form $H_n =$ "between n and $n+1\%$ of all people are brown-eyed," with $0 \leq n \leq 99$, and let us assume a uniform distribution of the initial probabilities of these hypotheses. Then, the inductive probability of H_n , given a sample of say m people, including $k \leq m$ with brown eyes, can be calculated based on Kolmogorov's probability axioms and the statistical principal principle ([sec. 4.7.2](#) – [4.7.3](#)). In summary, Popper's criticism of logical induction for empirical law hypotheses is unjustified or at least exaggerated; but it is certainly justified in the case of scientific theories.

Based on Popper's criticism of logical induction, we can evaluate competing theories only *comparatively* , that is, in relation to each other. Once, we have carried out an empirical success comparison of currently available alternative theories, we will base our future predictions and actions on those theories which up to now have been *most successful* , in other words, on the ones which have been corroborated best. Popper, of course, also sees it this way (1983, 65; 1979, sec. I.9). It follows from this that Popper's "deductivist" program of corroboration, too, contains a fundamental inductive step – a third kind of induction, that has been called epistemic induction ([Musgrave 2002](#)) or meta-induction ([Schurz 2008b](#)):

3 Epistemic induction or meta-induction says the following: If a theory T_1 has been more successful than T_2 *so far* (explanatory and prognostically), then it is probable, *relative* to the given state of evidence, that T_1 will also be more successful than T_2 *in future* . In other words, the success preferences established *so far* are projected inductively into the future. This induction principle is called "epistemic" or "meta-," because it is not object-level hypotheses (about ordinary events), but epistemic meta-hypotheses about the degree of corroboration of our object-hypotheses, which are being inductively projected. The epistemic induction principle is indeed indispensable for all empirical disciplines. Were this principle not to be accepted, there wouldn't be any point in Popper's method of testing. Success so far would then simply be irrelevant to our future actions. Although for example the view that heavy bodies on Earth are disposed to fall to the ground and not to float freely has been better corroborated so far than the opposite view, this would not be a reason to make this view the basis of our future actions, since logically, all bodies might just as well start floating as of tomorrow. Unfortunately, Popper kept an exaggerated distance to logical empiricism throughout his life,

and condemned any form of induction whatsoever. But his arguments on this topic in (1983, e.g. 66f) seem to rest more on mere terminological maneuvers than on actual material reasons (see [Schurz 2002a](#)). Other critical rationalists like Watkins ([1984](#) , 340f) and Musgrave ([2002](#)) have accepted the epistemic induction principle.

Epistemic induction, let us remember, is always only *comparative* . An important consequence of this is the following: according to epistemic induction, the degree of confirmation of a theory is always *doubly* relative:

- relative to the given state of observational knowledge, and
- relative to the current state of alternative theories.

2.7.3 Abduction and Inference to the Best Explanation

Scientific theories differ from empirical law hypotheses in that they contain theoretical concepts which speak about unobservable entities or properties. Inductive generalization enables us to infer from particular observations to empirical law hypotheses, but not from empirical law hypotheses to scientific theories. This is because inductive inferences cannot introduce new (theoretical) concepts into the conclusion, which are not contained in the premises. There is, however, one further – but controversial – kind of inference, by which this is possible, namely abduction, or inference to the best explanation. Roughly expressed, this is an inference from an observed *effect* to a likely *cause* , for example, from a curved trail in the sand to the recent passage of a sand viper.

At this point, we must warn against a *terminological misunderstanding* . A number of authors understand the concept of induction in a *broad* sense – they identify inductive inferences with *any* kind of content-expanding or nondeductive inferences, and so regard abduction as a form of induction (e.g. [Pollock 1986](#) , 42; [Bird 1998](#) , 13; [Ladyman 1998](#) , 28). By contrast, for the sake of a clear demarcation between distinct concepts, we understand induction always in the above outlined *Humean* sense, and see abduction as a second nondeductive or content-expanding kind of inference.

The abductive form of inference goes back to C.S. Peirce. In (1878a), Peirce explicated abductive inference formally as a retrodictive inference from “all Fs are Gs” and “this is a G” to “this is an F.” While this characterization of abduction does dominate the computer science literature on abduction (see [Flach/Kakas 2000](#)), it is much too narrow to grasp what the late Peirce was trying to achieve. The essential thing for Peirce ([1903](#) , §170) was that abductive inference allows new theoretical concepts and models to be introduced in the sciences. Newton, for example, inferred abductively from the movement of the planets around the Sun to the existence of a gravitational force. The epistemological inference from subject-internal perceptual experience to external empirical facts is also an abductive inference: the assumption that my visual experiences are being caused by objects existing independently of me is the best explanation for the constantly recurring regularities in my visual experiences (cf. [Schurz 2024 a](#), ch. 9). As is shown in Schurz ([2008a](#)), abduction comprises a whole *family* of kinds of inferences, which have the following pattern in common (cf. [Peirce 1903](#) , §189; [Niiniluoto 1999](#) , S439):

(Prop. 2.7-1) *General pattern of abduction* :

Premise 1 : A (singular or general) fact E, in need of explanation.

“ *Premise* ” 2 : A background knowledge K, which implies for a hypothesis H, that H is a potential and sufficiently plausible explanation for E.

Abductive conjecture : H is true.

As Peirce himself emphasized, the trustworthiness of an abductively acquired hypothesis is rather preliminary: the abducted hypothesis must go on to be empirically tested by deduction and induction in order to acquire the character of a probable hypothesis (1903, §171). More importantly, the given background knowledge always contains *several* possible hypotheses, H_1 , ..., H_n , which *potentially* explain the given evidence E (i.e. if they were true, they would explain E); and the abductive inference selects the most

plausible hypothesis among them. In this sense, Harman ([1965](#)) transformed Peirce's concept of abduction into the following inference to the best or best available explanation (IBE):

(Prop. 2.7-2) *Inference to the best (available) explanation – IBE* : Infer abductively to the most plausible explanatory hypothesis among all explanatory hypotheses for the given explanandum E that are possible, or are available in the given background knowledge, respectively.

Since the later Peirce also regarded abduction as IBE (1931ff, CP 6.525), abduction and IBE can be seen as approximately equivalent forms of inference (cf. [Douven 2011](#) ; [Williamson 2016](#) ; [Niiniluoto 2018](#) , 17f); however, this equivalence thesis is controversial (cf. the discussion in [Schurz 2023c](#) , ch. 1).

According to the understanding of Harman's ([1965](#)) and other representatives of IBE ([Lipton 1991](#) ; [Armstrong 1983](#)), an IBE infers from the explanatory success of a theory to its *likely truth* in the realist sense. Lipton ([1991](#) , 58) and van Fraassen (1989, 142ff) have pointed out that one can infer only to the best *available* explanation, because no one knows all possible explanations. In this situation, Kuipers ([2000](#) , sec. 7.5.3) has made an interesting suggestion. He brings the same objection into play as Popper made against logical induction: De facto, we can only ever reach *comparative* judgments, according to which a theory is explanatorily more successful and, hence, likely to be closer to the truth than all alternative theories *known* to us. Since this explanatory success includes also predictive success, Kuipers speaks more generally of "empirical success." Kuipers therefore suggests modifying the pattern of abduction in the following way: first, instead of speaking of an "inference to the best explanation," he speaks of an "inference to the best theory," and second, that a theory T is empirically most adequate among all known alternative theories does not lead to the conclusion that T is likely to be true, but only that T is *closest to the truth* among all known alternative theories. The interplay of the epistemic inductive inference and the abductive inference to the best theory is shown in [fig. 2.7-2](#) .

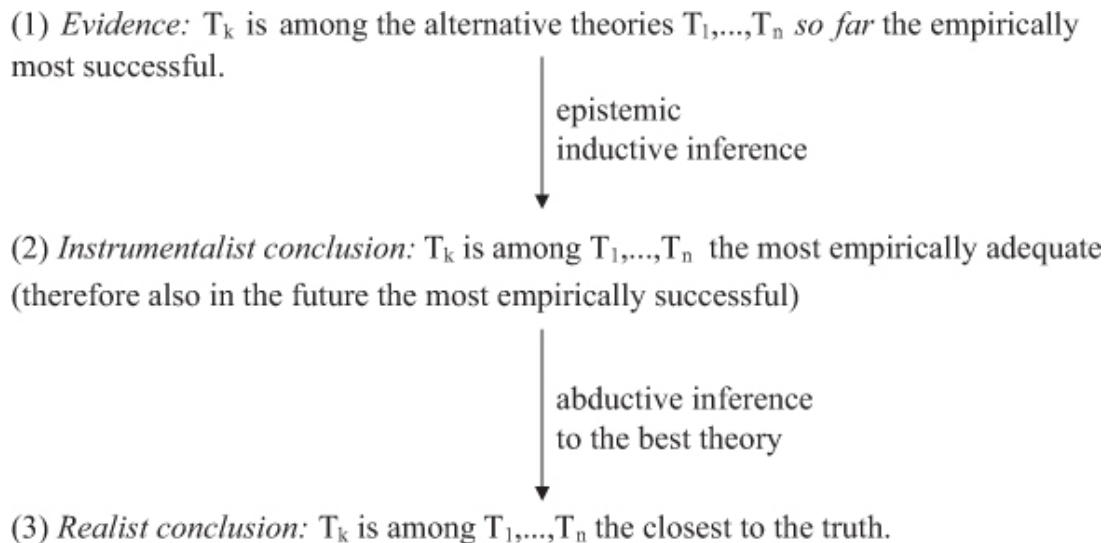


Figure 2.7-2 Interaction of epistemic induction and abduction. ↵

In [fig. 2.7-2](#) it becomes clear how these two kinds of inference are connected with the positions of *instrumentalism* and *realism* in the philosophy of science (see [sec. 5.7](#)). Instrumentalists, like for example van Fraassen, accept the inference from (1) to (2), but reject the abductive inference from (2) to (3), because according to them, theories may only be justified to be more or less empirically adequate, but not to be true or false in the realist sense (van Fraassen 1980, 11f, 20; 1989, 142f). Realists in the philosophy of science believe on the other hand in the likely truth or truthlikeness of well-confirmed theories; for them, it is that very inference from (2) to (3) which is essential.

2.7.4 Monotonic and Non-monotonic Inferences

Because deductive inferences are certain, that is, preserve truth (from premises to conclusion) in all possible worlds, they are monotonic in the following sense:

(Def. 2.7-1) *Monotonicity* : A valid inference (from premises $P_1, \dots, P_n$ to a conclusion C) is called monotonic, iff it remains valid after the addition of arbitrary further premises.

By contrast, inductive or abductive inferences don't possess the monotonicity property, because they are uncertain, that is, they don't transfer truth in all, but only in sufficiently "normal" worlds. We restrict the notion of *validity* to deductive inferences (indicated by a single forward slash "/"), while in the case of uncertain inferences, we speak instead of their (nondeductive) *correctness* (indicated by a double forward slash "//"). With this notation, monotonicity means that if " $P_1, \dots, P_n / C$ " is valid, then " $P_1, \dots, P_n, Q / C$ " is valid for every new premise Q .

The non-monotonicity of uncertain inferences makes it possible that " $P_1, \dots, P_n // C$ " is correct, but " $Q, P_1, \dots, P_n // C$ " is incorrect for some (true) premise Q , where Q typically expresses "exceptional information." For example, "1000 swans have been observed to be white // Therefore probably: All swans are white" is a correct inductive generalization, but the premise-expanded inference "1000 swans have been observed to be white and 10 to be black // Therefore probably: All swans are white" is inductively incorrect. The logical examination of *non-monotonic* inferences has become a significant area of inquiry since the 1980s (cf. [Gabbay et al. 1994](#) ; [Schurz 2004](#)).

A non-monotonic inference with true premises has only *preliminary* value and can be made worthless by new (true) information. In epistemology, non-monotonic inferences are therefore also referred to as *defeasible* (cf. [Pollock 1974](#) , 36ff). The preliminary nature of non-monotonic inferences is expressed by the fact that they have a different *epistemic application rule* than monotonic inferences. Here, we take the *belief system* (or 'knowledge' system) S of a person or group of persons to be the set of propositions which are justifiably accepted as true by these persons.

(Prop. 2.7-3) *Application rule for monotonic inferences* : If a deductive inference is valid, then you should add its conclusion to your belief system S , as long as S contains the premises.

(Prop. 2.7-4) *Application rule for non-monotonic inferences* : If an uncertain inference is correct, then you should add its conclusion to your belief system S , as long as S contains the premises, and these premises cover *all* propositions in S which are of relevance to the conclusion. For inductive-probabilistic inferences, [prop. 2.7-4](#) is equivalent to Reichenbach's rule of the narrowest reference class and Carnap's requirement of total evidence (see [sec. 3.9](#) , [3.13.4](#) , [6.3.2](#)).

Complementary and Advanced Topics

2.8 How Much Metaphysics is Presupposed by Science?

2.8.1 Truth, Idealism, and Realism

One may object that our reconstruction of the supreme goal (G) of the sciences and the assumption ($E1$) of minimal realism assumes too much of realism. Famous scientists and philosophers of science, such as George Berkeley (1685–1753), Ernst Mach (1836–1916), and the early Rudolf Carnap, were idealists or positivists, that is, anti-realists. In their view, science is possible if it is merely based on human intersubjective experiences, *without* the presupposition of a subject-independent reality that transcends these experiences. We do not deny this possibility.

Admittedly, our assumption ($E1$) makes a (weak) realistic assumption, which is discussed below. Some philosophers argue that already our explication of the goal (G) of science, the discovery of content-rich

truths, presupposes realism, insofar as truth means correspondence of statements with an assumed reality. But this is not the case, since (G) does not say that the domain of phenomena to which true statements correspond has to belong to an external and subject-independent reality, as claimed by realists. The domain of phenomena called ‘reality’ may also consist of experiences, as assumed by idealists.

The major epistemological positions can be summarized as follows (see [Dancy 1985](#) , 136ff; cf. [Schurz 2024a](#) , [sec. 2.7](#)). The domain of entities to which true statements correspond consists either of

- a. one’s own subjective experiences (*subjective idealism* or solipsism),
- b. the experiences of all subjects (*intersubjective idealism*),
- c. the *possible* experiences of all subjects (*possibilistic idealism*), or
- d. a subject-independent reality that causes the experiences (*realism*).

Solipsism is a rather extreme position (most idealists have been possibilistic idealists). Be that as it may, the important point to observe here is that the concept of *truth* makes sense in *all* of these positions. Hence, truth in the sense of correspondence theory is not realistically biased but *metaphysically neutral* (cf. [Kirkham 1992](#) , [sec. 6.5](#)). So also, our formulation of the supreme goal (G) of science is neutral. It does not entail that successful science presupposes the epistemology of realism. The goal of truth makes sense for *all* of the epistemological positions; only the *kinds* of truths that, according to these positions, are pursued by science are different (true descriptions of experiences versus true descriptions of subject-independent facts).

While (G) is metaphysically neutral, our epistemic assumption (E1) is not. It asserts that realism, in the explained weak version of it, is the *normal* epistemological assumption of science. I think this claim is defensible: more or less all empirical scientists assume the external reality of their objects of investigation; idealist or anti-realist worldviews are the clear exception in normal science, although the realistic attitude of normal science can be rattled in scientific revolutions (recall [sec. 2.1](#)).

What distinguishes our minimal realism from all sorts of *metaphysical realism* is that neither the existence of a subject-independent reality, nor its knowability, is assumed to hold on a priori reasons or is a necessary presupposition of knowledge. By contrast, the minimal realism of assumption E1 is a *hypothetical-constructive* realism, insofar as we consider realism not as a presupposition, but as a hypothesis or model construction – in analogy to scientific hypotheses. Our realism is committed to *methodological naturalism* , as the hypothesis of realism must be corroborated by the *empirical success* of our cognitive belief system, as the *best explanation* of this success. This abductive form of realism corresponds at the same time to a variant of empiricism, namely *abductive empiricism* , which – unlike the reductionist empiricism of the neopositivists – is compatible with realism ([Schurz 2024a](#) , 90).

Because of its naturalistic tenets, hypothetical-constructive realism is compatible with legitimate doubts about a ‘full’ realism found in the philosophical interpretations of *quantum mechanics* . In all of its interpretations, quantum mechanics assumes the existence of a microphysical reality that possesses some subject-independent and discoverable properties (so-called observables). Only certain properties of entangled (or superposition) states are not objectively measurable, because it is a postulate of quantum mechanics that every measurement transforms an entangled state into a statistical ensemble of pure states (see [Herbert 1985](#) , sec. 8+9; [Penrose 2007](#) , sec. 29).

Hypothetical-constructive realism is also compatible with *structural* realism. In its *epistemic* version, structural realism says that only the structure but not the intrinsic nature of the entities making up reality are knowable ([Worrall 1989](#) /1997; [Psillos 1995](#) ; [Votsis 2011](#)). The *ontic* version of structural realism is motivated by quantum mechanics and claims that even the entities that make up reality are themselves not individuals but structures ([Ladyman 1998](#) ; [French/Ladyman 2003](#)).

2.8.2 Constructive Realism versus Radical Constructivism

Hypothetical-constructive realism differs markedly from radical *constructivism* mentioned in [sec. 1.3.9](#) . The latter position was developed by psychologists such as Glasersfeld ([1995](#)) or biologists like Maturana and Varela ([1992](#)). Historically speaking, radical constructivism is a modern variant of subjective idealism. The main argumentation of this philosophical position is based on the fact that our perceptions and concepts of reality are cognitive *constructions* of our mind or brain, respectively. Radical constructivists, however, infer from this hardly deniable premise that even ‘reality’ itself is a construction, that is, a reality that is

independent of our constructions does not exist. A similar argument has been put forward by social constructivists such as Latour and Woolgar ([1986](#)), in regard to the ‘construction’ of scientific facts by scientists.

In this subsection, we want to show that the argument of radical constructivists is based on a *fallacy*. We reconstruct this argument as follows (e.g. [Glaserfeld 1995](#), 40, 51; [Kriz et al. 1990](#), 19f):

(2.8-1) *Premise of radical constructivism = epistemic constructivism* : Our perception and mental representation of reality is not something given, but the result of an active cognitive construction.

Conclusion of radical constructivism = ontological constructivism (idealism) : Therefore reality, too, is not given “in itself,” but “constructed” or “constituted” by us.

The crucial fallacy of this argument is the ‘magical transformation’ of epistemic into ontological constructivism (cf. [Sokal/Bricmont 1998](#), 102). Hypothetical-constructive realism shares the premise of ([2.8-1](#)) with radical constructivism, but it definitively rejects the conclusion. The *invalid* step from epistemic to ontological constructivism is frequently substantiated by constructivists with the help of the following ‘justification’ (cf. [Kriz et al. 1990](#), 20; [Glaserfeld 1995](#), 51), which may be called the *Kantian fallacy*, because it corresponds to a popular interpretation of Kant’s philosophy:

(2.8-2) “*Justification*” of radical constructivism : It is impossible to know something about reality *in itself* – independent of its epistemological construction. The only thing about which we can know something is the reality *constructed by us*.

The subtle error of this justification attempt is that the concept of “knowing” is understood in the *naïve-realist* sense of a kind of direct mirror-imaging. Then it is indeed impossible to know something about an object that is “independent of our epistemological construction” – which in this case means: “*not influenced* by this construction.” Modern truth theory, though, understands the relationship between a true statement and its object not as a quasi-identical mirror-image, but as a *structural correspondence*, which *conveys certain information* and has to be neither complete nor unique. In this view, there is no incoherence in saying that *by means of* our epistemological constructions we know something about a reality which exists *independently* of our epistemological constructions. For the independently existing reality is not taken to be immediately given to our mind, as in metaphysical or ‘direct’ realism, but is postulated as a system of hypothetical entities, whose existence can best explain the factual success of our knowledge. Giere ([1999](#), 174ff) and Kuipers ([2000](#), sec. 13) have called this position “constructive realism,” and we call it *hypothetico-constructive* realism or alternatively *abductive* realism, in order to emphasize that realism too is ultimately a fallible hypothesis.

In conclusion, it is the very argumentation of radical constructivists which clings to naïve-realism, inasmuch as it uses a naïve-realist interpretation of the objects of our knowledge to prove that these objects are subjectively constructed; while from the hypothetical-realist view, the statement ([2.8-2](#)) and so the whole argumentation of radical constructivists is untenable.

2.9 On the Theory (In)dependence of Observation

Classical empiricists argued that all scientific (or cognitively valuable) concepts are reducible to observational concepts. That this is not so was convincingly demonstrated by the recognition of *theoretical concepts* in the middle and later stages of logical empiricism (cf. [Carnap 1936](#) /37, 1956a). Theoretical concepts go beyond what is observable – this is a generally accepted insight in post-positivistic philosophy of science ([sec. 5.1](#)). Accordingly, the empiricism in our reconstruction of the common epistemological ground of science ([sec. 2.2](#)) is reduced to a minimalistic version according to which theories, although going beyond what is observable, must at least have *some* observable consequences by which they are testable.

What minimalistic empiricism still assumes is the existence of at least *some* theory-independent observation concepts and sentences, which can be used to independently test competing theories. However, even this minimalistic empiricism came under attack in the post-positivistic philosophy of science. The thesis of the *theory-dependence* of all observation and experience has been made popular by Hanson ([1958](#)), Kuhn ([1962](#)), and Feyerabend ([1975](#)); with precursors in Wittgenstein ([1945](#) /2001, part 2, §xi) and Popper (

[1935](#) /2002, new appendix X). Meanwhile, this thesis seems to be accepted by the majority of philosophers of science. Closer inspection reveals, however, that what is accepted by the majority of philosophers is a cautious and loose version of this thesis, that leaves plenty of room for interpretations that are compatible with our epistemological assumptions (E3) of objectivity and intersubjectivity, and (E4) of minimal empiricism. For many philosophers of science, what is important in the theory-ladenness thesis is a critical attitude toward what is called “data” in science: these data are typically theory-laden interpretations of the scientist’s perceptions, and the assumptions made in these interpretations are often concealed behind a veil of expertocracy. Nevertheless, I will try to show in this section that in all empirical sciences, one can find theory-neutral observation concepts and statements that scientists may draw on in cases of conflicts of opinion.

Before I start my enterprise, let me try to demonstrate the significance of this question for the philosophy of science. A *complete* theory-dependence of observations, in the sense that *all* observations are theory-dependent, would turn empirical science into a *circular* affair without any hope of objectivity (cf. [Nagel 1979](#) , 79). For suppose that two adherents of rival theories T_1 and T_2 are observing the same area of space-time, and assume that their respective theories imply the contradictory predictions P_1 and P_2 . If complete theory-dependence were given, then the observations of the theory-adherents would be determined by their theoretical *expectations*. The supporter of T_1 would then observe P_1 where the supporter of T_2 observes P_2 . So, each prediction would be a *self-fulfilling* prediction, and an independent empirical test of theories would be impossible.

Complete justificatory circles are demonstrably epistemically worthless (see [sec. 2.11](#)). By contrast, a *partial* theory-dependence, in the sense that there are at least some theory-neutral observations on which theory-dependent scientific data rest, is not a threat to the possibility of objective empirical tests. Adam ([2004](#)) and Brown ([1993](#)) have studied cases in which the recognition of a scientific datum D that falsifies a theory T relies on the assumption that theory T is true. This sounds paradoxical, but in fact is not: the case is harmless, because the datum D was gained from a theory-neutral observation O using T . This means (in the simplest case) that T entails the implication *if O , then D* . Moreover, that T is falsified by D means that D entails not- T . Taken together, this implies by simple logical means that O entails not- T , that is, that the observation O falsifies T . Nothing paradoxical is involved in this situation.

In conclusion, for a defense of the objectivity of empirical science, it is sufficient to refute the thesis of complete theory-dependence, that is, to establish the existence of *some* theory-neutral observation concepts and statements in science. In the next subsection, I will sketch eight central arguments in favor of different forms of theory-dependence as well as the corresponding counterarguments. I will try to show that although the pro arguments make correct points, they do not exclude the existence of theory-neutral observations in the narrow sense of perceptions.

2.9.1 Theory-Dependency of Observation – Pros and Cons

Pro-1 : Observations are expressed as realist statements . Observation statements such as “this thing is green” are usually expressed in the realistic mode. So, they imply the *hypothesis of realism* , that is, the existence of subject-independently existing objects (“this thing”) that cause our subjective perception (cf. [Popper 1935](#) /2002, sec. V, §25, for a version of this argument). Observation sentences in this realist interpretation transcend the content of subjective experience. We can be in *error* about them, for example, because our sense experience may rest on a *hallucination* .

Contra-1 : This argument is true, but this is not what is meant with theory-dependency of observation. What is meant is that the *content* of one’s observations depends on the *content* of one’s background theory (expectation or worldview). But the question whether one is an epistemological realist or a phenomenalist idealist does not affect the content of one’s observations. It merely changes the interpretation of the term “thing:” in the realist interpretation, this term refers to assumed entities outside of one’s experiences, and in the phenomenalist interpretation, it refers to entities within one’s experiences. But whether one is a realist or an idealist does not turn a green-observation into a blue-observation (this argument can be found in [Carnap 1928b](#)).

Pro-2 : Observations are theory-guided . The selection of relevant problems and experiments is essential in every scientific research, for reasons of complexity alone. This selection is guided (among other things) by one's theory: which theory I support determines which observations I look for. ³

Contra-2 : It is true that the selection of one's observations depends on one's theories, and more generally, on one's interests. But that observations are selected in a theory-dependent way does not imply that the observations *themselves* are theory-dependent, in the sense that supporters of different theories would observe different things when perceiving the same spatiotemporal region or measurement result. Supporters of the theory T_2 *could* also make the observations made by supporters of the theory T_1 and *would* then recognize their truth. More generally speaking, the fact that empirical research is interest-driven does not imply that certain observations *cannot* be made, but at most that these observations are not *wanted* to be made, or that they are *ignored* . However, such a dogmatism is a sign of *bad* science and violates the principle of *total evidence* ([sec. 3.13.4](#)).

Pro-3 : Perceptions are unconscious theoretical constructions. That perceptions or observations in the *narrow* sense are unconscious theory-like constructions is an insight that goes back to Helmholtz ([1896/1910](#)). The brain's visual regions use refined neural construction mechanisms to create a consciously experienced 3D image from the unconscious 2D retinal light-impressions. Normally, this construction process generates a unique 3D image within milliseconds, so that its constructive character is *hidden* from the consciousness: you think you see "reality itself" (for this reason, humans are intuitively naïve-realists). But in certain – typically in artificially engineered – situations, this visual construction process reveals its constructive character, because it leads to consciously experienced visual *ambiguities* , *illusions* , or other strange effects. This has been impressively demonstrated by the findings of cognitive psychology (cf. [Rock 1984](#)). Simple examples of ambiguities are *tilting pictures* such as Jastrow's duck-rabbit, a picture that can be seen either as a duck or a rabbit (see [fig. 2.9-1](#)). The duck-rabbit has been used by Wittgenstein ([1945/2001](#), part II, §xi), Hanson ([1958](#) , 12), and Kuhn ([1962](#) , 112ff) as evidence for the 'theory-dependence' of perception, insofar as what we *expect* to see influences what we see.



Figure 2.9-1 Duck-rabbit. [↗](#)



Figure 2.9-2 Mueller-Lyer illusion [↗](#)

An example of an *optical illusion* is the Mueller-Lyer illusion in [fig. 2.9-2](#) . The explanation of this illusion lies in mechanisms of 3D-vision: the arrows and inverted arrows produce the 3D-effect that the right stroke appears to be closer to the observer than the left stroke and is therefore perceived as being shorter, although both strokes have equal length (see [Rock 1984](#) , 163, 167).

Contra-3 : Neither the logical empiricists nor Popper had good replies to the psychological arguments concerning the constructive character of perception, because according to their position of "logical purity," philosophy should be independent from psychology. What counts as an observation sentence was logically speaking a matter of "pure convention" (cf. [Popper 1935/2002](#), sec. 5, §30; [Neurath 1934](#)). We now present some arguments against the theory-ladenness of perceptions that are themselves based on findings from cognitive psychology.

The findings about the constructive character of visual processes refute only *direct realism* ([Dancy 1985](#) , 144). According to this epistemological position, we see things directly, as they are 'in themselves,' without

any mediating processes that construct intermediate images. In the light of contemporary knowledge about the neuro-cognitive process of vision, direct realism seems to be an untenable position (see [Schurz 2024a](#) , [sec. 3.5](#)).

However, all findings of cognitive psychology also confirm that our visual perception processes and their results are quite stubbornly *independent* of our *learned* information. Everyone experiences the right stroke in [fig. 2.9-2](#) as shorter than the left one, even if she has been informed that this visual experience is an illusion. These facts motivated Fodor's thesis of the modularity and impenetrability of vision ([Fodor 1984](#) , 34ff). Also, the Helmholtzian constructivist Rock ([1984](#) , 228) acknowledges the independence of vision from acquired information.

In the tilting picture of [fig. 2.9-1](#) , everybody is able to perceive it either as a duck or as rabbit and to switch back and forth between these two gestalts. Kuhn argued that a similar gestalt switch took place when astronomers saw Uranus before 1781 as a fixed star, but after 1781 as a moving planet (1962, 111, 121f). But it does not seem right to conceive the Uranus case as perceptual switch: The visual content of the Uranus-perception is always a bright spot in the sky, which is perceptively stable and does not switch. The question whether it represents a fixed star or a planet is not a question of perception but of theoretical interpretation.

The processes of human vision are based on innate mechanisms that have developed in millions of years of evolutionary history. Their selection history explains why the *same* mechanisms, which under normal environmental conditions give a correct visual depiction of reality, can create optical illusions under certain non-normal conditions. Evolutionary considerations also explain why human perception, in order to be evolutionary successful, must be independent from expectations, for otherwise it could not correct *wrong* expectations. As an example, consider prehistoric men entering a cave which is inhabited by a cave bear that they did not expect: if their perception would not immediately have corrected their wrong expectation, these prehistoric men would not have survived.

By "theory-dependence" of perception, one could at best understand that perception is dependent not on learned information, but on 'innate theories' which all humans share, because they are rigidly wired up in their visual cortex. However, we do not want to understand the concept of "theory" in such an overstretched way. By the theory-dependence of perception, we always only mean the dependence of perception on *acquired* information that may vary in different subjects, cultures, or scientific communities. It seems that visual perception in this sense is – at least to a large extent – theory-independent.

Pro-4 : Sometimes, visual perception depends on one's pre-expectations . The last statement of Contra-3, that visual perception is independent from pre-expectation, is not generally true – it holds only for clearly presented visual stimuli. There exist some psychologically well-investigated experiments, to be discussed below, in which visual perception does indeed depend on pre-expectations.

Contra-4 : This is only the case for the stimuli at the perceptibility threshold , that is, when the limits of perceptual recognition capacities are reached (cf. [Brewer/Lambert 2001](#) , S181). These limits are reached in three kinds of cases:

1. *Poorness* of the stimulus: For example, (1.1) in a dark forest at night, anxious persons seem to see the contours of mysterious figures in the dark thicket, which disappear at daylight. Or (1.2), if the relevant stimulus is too small to be seen, for example, when you ask which of two approximately equally long strokes is the longer one, the test persons' answers may become expectation-dependent. The perceptibility threshold is also reached when (1.3) the presentation time of the stimulus is very short, as in high-speed reading tasks (cf. [Estany 2001](#) , 208).
2. *Ambiguity* of a stimulus that has no unique gestalt solution, such as the duck-rabbit in [fig. 2.9-1](#) : Here, one's expectation about the gestalt influences which gestalt one sees first (though one may switch the gestalt at will).
3. A *high complexity* of the relevant stimuli patterns is a third reason, not for expectation-dependence, but for dependence on acquired knowledge: here, visual *training* may be necessary to detect the relevant visual patterns. As an example, think of the radiologist who recognizes pneumonia in the X-ray image of your chest.

Because of these finding, some cognitive psychologists described human vision as a combination of bottom-up processes (information flow from the retinal cells) and top-down processes (information flow from higher

cognitive regions; cf. [Estany 2001](#) , sec. 4). Although this description is not false, it misses the important point of this combination, namely the *priority ordering* in favor of bottom-up information. Top-down information from expectation by memory can correct visual bottom-up information *only* if this bottom-up information is *too poor* to allow a clear gestalt interpretation. Otherwise, bottom-up information *overrides* top-down expectations. This is illustrated, for example, by an experiment in which a test person looks through a peephole with one eye at two white balls of equal size and distance in an otherwise dark black box ([Rock 1984](#) , 80f). In this experiment, all cues about the distance of the two balls coming from bottom-up information have been eliminated: there is neither (a) binocular depth information nor (b) monocular depth vision (e.g. by motion parallax via eye movement). If one tells the test person that one is a tennis ball and the other a ping pong ball, the tennis ball is perceived as being further away. This illustrates the utilization of acquired information when bottom-up information is missing. However, as soon as cues of sort (a) or (b) are admitted, for example, when the peephole is broadened, and the person is allowed to move her eye or to look with two eyes, the illusion disappears, and the person sees the two balls in equal size and distance.

Pro-5 : Scientific data (or observations i.w.s.) are theory-dependent. Scientists commonly use a wider concept of ‘observation’ or ‘datum’ than perception with the ‘naked eye.’ According to the scientific usage of this word, scientists observe galaxies, bacteria, or electrons with suitable *instruments*. A number of authors, including Kuhn in many passages, relate their thesis of theory-dependence to this observation concept i.w.s. [Giare \(1999 , 12f, 80\)](#) and [Kuipers \(2000 , sec. 13.3\)](#) argued that whenever a new measurement instrument is invented, a formerly theoretical concept is turned into an observational concept.

Contra-5 : Scientific data are theory-dependent interpretations of perceptions. The theories presupposed here are mainly theories about the way measurement instruments work, which are usually treated as an unproblematic part of the advanced knowledge of a scientist. This *pragmatically* extended observation concept i.w.s. is undoubtedly very important for the efficient progress of science, since one cannot always put everything simultaneously into question, and the investigation of scientific problems rests on large amounts of background knowledge that is highly corroborated and taken for granted. However, this fact does not imply what some philosophers of science wrongly infer from it, namely that there *are* no observations in the narrow sense (i.n.s.) in science, that is, direct *perceptions* (mainly visual perceptions). We claim theory-independence only for this narrow concept of observation.

Most of the time, scientists agree on their observations i.w.s., and there is no need for them to make their direct perceptions explicit. But when scientists disagree on what they have observed i.w.s., they will critically reflect their implicit theoretical assumptions that led them to infer different observations i.w.s. from their common perceptions. For example, a pulmonologist A may report to have seen pneumonia in the X-ray image of a patient’s chest. What he has seen directly is a certain blackening pattern that he interprets as pneumonia. If another pulmonologist B doubts this diagnosis, then what B doubts are not the visual abilities of A. Rather, A and B will discuss whether the blackening pattern which they *both* see is indicative of pneumonia or not. In conclusion, the epistemological significance of observations i.n.s. lies not in being permanently verbally reflected, but in serving as a possible level of theory-neutral agreement to which scientists can step back in the case of disagreement about observations i.w.s.

Pro-6 : There is a continuity between observational and theoretical concepts. Maxwell ([1962](#) , 7), Carnap ([1966](#) , 255f), and Hempel ([1966](#) , 78–84) pointed out that there is a continuous transition from observability with the naked eye to intermediate stages such as observability with glasses or a magnifying glass, until observability with an optical telescope or microscope, an electron microscope, or a radar screen. A number of authors inferred from this that any demarcation between perception and theory-dependent observation i.w.s. is arbitrary.

Contra-6 : That there exists a continuous transition in the extent to which concepts are theory-laden is true. But first , the fact that there exists a continuous transition between two features like for example black and white, does not by any means imply that there is no difference between black and white (cf. van Fraassen 1980, 214). So, the distinction between observation in the narrow and in the wide sense is significant even if it is continuous. *Second* , there are marked jumps in this transition: while observation i.n.s. does not require any learned background knowledge, the required background knowledge increases successively for the usage of the above-named instruments. As an illustration of this point, imagine what a native tribesman would do with these instruments, without having received instructions. Presumably, he would understand glasses, but

already optical microscopes and telescopes would hardly be understandable, and that's not to mention electronic microscopes or radar screens, as the tribesman has no clue about the background knowledge that is needed for an adequate handling of these instruments.

Pro-7 : Meaning holism makes observation concepts theory-dependent. Quine ([1951](#)) doubted the traditional analytic-synthetic distinction (see [sec. 3.11](#)). His alternative semantic view was a radical meaning holism, according to which the meaning of a concept is determined by all that we know about it. If this were true, then theory-neutral observation concepts could not exist, because in scientific theories, observation concepts get connected with theoretical concepts. According to meaning holism, this would imply that their meaning becomes loaded with theoretical statements. In this sense, Chalmers ([1986](#) , 36) argued that “water” is a theory-laden concept, because chemical theory connects it with the molecular structure H_2O .

Contra-7 : Radical meaning holism will be criticized in [sec. 3.11](#) as more or less *untenable* . Quine's inner conflict concerning observation concepts is revealed in (1992, 7f), where he attributes a “two-faced nature” to the sentence “this is water:” “holophrastically,” so Quine, the sentence is free of theory, but by its composite structure, the sentence contains theory. This sounds like an ad hoc maneuver to escape the untenable consequences of meaning holism. We argue that the meaning of observational concepts is fixed by processes of definition or ostension such as “this is water.” This meaning is *not* affected by theories, which connect observational with theoretical concepts. If this were not so, then one would have to say that a physicist and a layperson mean entirely different things when they point to a creek and say that water is bubbling there – but this seems too strange to be true.

Pro-8 : Observation concepts (as all other concepts) are language-dependent. The thesis of the language dependence of experience was made popular by the philosopher Wilhelm von Humboldt and by the anthropological and ethnological research of Sapir and Whorf (see [Kutschera 1976](#) , [sec. 4.1](#)). Many native tribes, for example, have different and less differentiated concepts of color than members of Western civilization ([Berlin/Kay 1999](#)). From this fact, cultural anthropologists inferred that these natives really see the colors in their environment differently from us. The other way around, Inuits (Eskimos) have reportedly many more concepts for sorts of snow than us (see [Whorf 1956](#) , 216; [Kutschera 1976](#) , 251f). Whorf derived from this his *linguistic relativity principle* , according to which we can only perceive what is describable, and in a sense also prescribed, by the concepts in our language.

Contra-8 : It must be admitted that different cultures, to a certain degree, develop different systems of concepts, in which they describe the perceptions of their different environments. But from this, it does not follow that their sense experience itself is language-dependent. We could only conclude this if members of these cultures were also not able to *learn* the observation concepts of other cultures by *ostensive training* , that is, by acts of pure ostension. This ostensive learning ability is present in all cases, however. The Zuni Indians can learn the concept of “orange,” for which they don't possess a word, through brief ostensive training – that is, through simple acts of pointing (“this is orange,” “this is not orange”). The Tiv and the Dani were able to learn many European color concepts in this way, for which there is no word in their mother tongue ([Garnham/Oakhill 1994](#) , 49–51; [Berlin/Kay 1999](#) , 24). Similarly, a European can learn the Inuit's words for different kinds of snow ostensively.

2.9.2 A Definition of Theory-Neutral Observation Concepts

In this section, we present our definition of theory-neutral observation concepts. We start from the insight that the defining criterion cannot come from logic, but must be an empirical-psychological one. Hence, the definition of “observation concept” will be relative to the sensorial capacities of humans. For example, the magnetic field is non-observable for humans, but observable for animals with magnetic sense organs (cf. [Carnap 1936](#) /1937, 454f; van Fraassen 1980, 56f).

Traditionally, an observational concept is characterized as a concept whose applicability to an observable object can be determined through mere perception (e.g. [Carnap 1936](#) /1937, 454f). To verify this characterization, we must rely on the *introspection* of individuals – but it is well known that many people indiscriminately conflate perception with theoretical interpretation in their supposed observations. For example, esotericists claim to observe “spiritual fields,” and 18th-century chemists reported observing

phlogiston. To avoid this problem, we base our definition on an experimentally testable criterion: the *ostensive learnability* of concepts. With this step, we do not only circumvent the problem of the unreliability of introspection, but also the influence of language. An *ostensive learning experiment* is based on nonverbal behavior and involves only three communicative acts that are culturally universal: acts of ostension (“this”), interrogative forms (“is this X?”), and words or gestures for “yes” and “no.”

It is essential for an *ostensive learning experiment* to work in a *nonverbal* way. Otherwise, the linguistic relativist could argue that the Zuni have learned to see the color “orange” only by learning the Western color language. Also, Quine says in several places that the characteristic property of observation sentences is their ostensive learnability, independently of acquired information (1960, 42; 1974, 37, 41f; 1992, 3f, 5f). But this thesis of Quine is in strong tension with Quine’s radical meaning holism (cf. pro-and-contra 8 of [sec. 2.9.1](#)). So, in spite of these similarities, Quine’s theory of observation sentences is essentially different from our account.

An ostensive learning experiment consists of a *training phase* and a *test phase*. In the training phase, the concept to be tested for observability is introduced to the test person (tp) with an *artificial word* “X” unknown to them, so that the tp does not associate *any* meaning with the artificial word. The concept referred to by “X” is illustrated to the tp with a set of positive and negative instances of the concept, and the information “this is an X” or “this is not an X” is given. The instances may be presented in the form of photos, movies, or real objects. They must be clearly visible by the tp; more precisely, the experiment takes place under ‘normal observation conditions’ (see below).

The ostensive learning process to be achieved by the tp consists in the abstraction of the meaning of “X,” that is, the property cluster corresponding to X, from the examples presented. Whether the tp has succeeded in this is investigated in the *test phase* which follows, in which a series of *new* positive and negative instances of the concept are presented, and the question “is this an X?” is asked. The tp’s success frequency in correct answers in the test phase indicates to which extent the tp has successfully abstracted from the training instances the perceptual property cluster that corresponds to the concept “X.”

If a concept can be learned from a small set of learning instances by (almost) all test persons, *independently* of their background culture and even if the tps possess no word for this concept in their spoken language(s), this is a more-or-less compelling reason to see this concept as a *theory-independent* observation concept. Without this assumption, the learning success would be a sheer miracle. If a Zuni Indian were unable to perceive the color feature “orange” because she has no concept of it, then she would not perceive the difference between the orange, red, and yellow pictures in the training phase, and it would be *inexplicable* how she could successfully pass the test phase. In conclusion, we propose the following definition of observation concept:

(Def. 2.9-1) A concept is a *theory-neutral observation concept* (or an observation concept i.n.s.) iff almost all humans can acquire this concept in an ostensive learning experiment, under normal observation conditions, independently of their background information, language, and culture.

Thereby, the *normal observation conditions* include (i) the normal physical conditions (sufficient light etc.), (ii) normal biological conditions (no biological deficiencies in the person’s sensory organs and nervous system), and (iii) normal psychological conditions (the person is in a wakeful and mentally normal state, and her yes/no utterances are truthful). All three conditions can be experimentally tested.

Our definition of observation concepts does not imply that all cultures have the same system of observation concepts (this is de facto false, as we know). What the definition implies is that different cultures can learn each other’s observational concepts in an ostensive manner, without being forced to adopt the language or worldview of the other culture. This consequence accords with the facts mentioned in [sec. 2.9.1](#) (contra-8), namely that native tribes can learn the Europeans’ color concepts, or that Europeans can learn the Inuit’s concepts for kinds of snow (etc.).

It would be interesting to perform ostensive learning experiments of the described sort (according to my knowledge, this has not been done so far). Generally speaking, the number of instances that are necessary to grasp the perceptual property cluster that corresponds to the concept “X” will not only depend on the question whether the property cluster is observable in principle or not, but also on the *complexity* of this property cluster. So, the result of an ostensive learning experiment is expected to be a *gradual learning curve*

. This turns the resulting notion of an observation concept into a *gradual* one: a concept may be more or less observational. Nevertheless, we expect that there will be clear statistical divide between the following types of concepts, as illustrated in [fig. 2.9-3](#) (cf. [Schurz 2015b](#)):

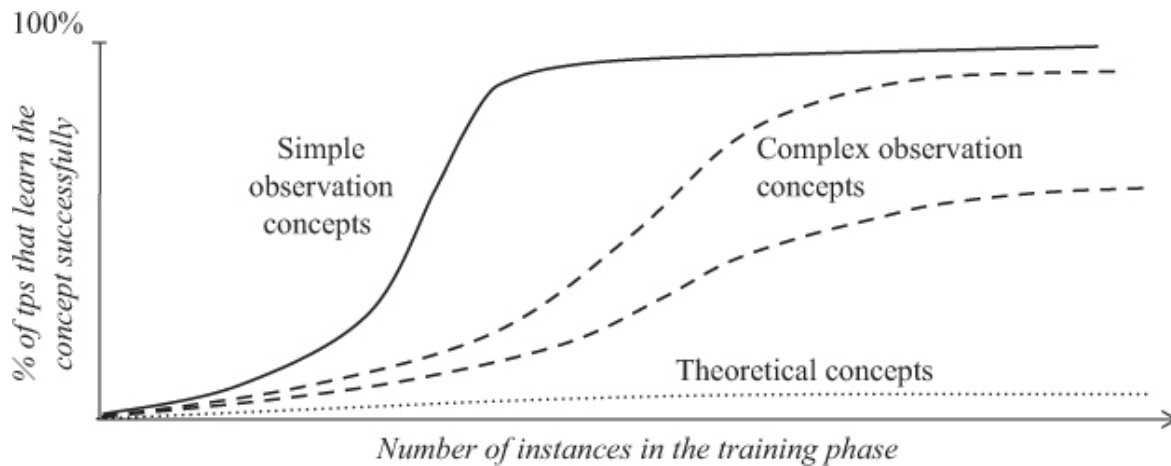


Figure 2.9-3 Expected learning curves in ostensive learning experiments. [↗](#)

1. For *simple* observation concepts (such as “red,” “square,” or “bird”), success frequencies will after a small number of instances climb above the significance threshold, toward values slightly smaller than 100% of tps.
2. For *complex* observation concepts (such as “either square and red, or circle and blue” or “hooded crow”), the success rate will climb slowly and will need many instances; often the success frequency will stagnate at lower levels, for example, at 50%.
3. For genuinely *theoretical* concepts (such as “atom,” “electric force,” or “oxidation”), the learning curve will not climb at all, even after many instances.

2.10 Value Neutrality in the Context of Contemporary Debates

2.10.1 The Role of Values in Science

As pointed out in [sec. 2.5.2](#), the value-neutrality requirement ([prop. 2.5-1](#)) possesses a critical-empiricist function. This stands in contrast to an influential division of epistemological paradigms according to their respective knowledge-guiding interests, which goes back to Habermas ([1966](#), 241ff) (cf., e.g., [Kriz et al. 1990](#), 151; [Rosenberg 2008](#), 127ff):

<i>Epistemological paradigm</i> :	<i>Knowledge interest:</i>
1 empirical-analytical	technical-instrumental controllability
2 hermeneutic	establishment of normative consensus
3 critical-dialectical paradigm	emancipation from power structures

Given our understanding of scientific value neutrality, these assignments are clearly incorrect. Science, being classified by Habermas as empirical-analytical, cannot be assigned to *any* specific interests that are external to science, but serves directly merely the science-internal epistemic goals, the search for content-rich truths. Scientific knowledge may be equally well used for technical purposes or for purposes of political reform. Adherence to value neutrality certainly does not mean that a scientist must refrain from practical involvement in society, or remain unaware of her social *responsibility* (cf. [Prim/Tilman 1979](#), 139f; [Koertge 2000](#), S48). All that the value-neutrality requirement implies is that the scientist *separates* her scientific knowledge from

fundamental value assumptions which she assumes in means-end inferences, and thereby *relativizes* her means-end inferences hypothetically to these value assumptions (recall the example in [sec. 2.5.2](#)).

The true difference between Habermas' three epistemological paradigms lies *not* in differing interests, but in the different treatment of the *value-neutrality question* (cf. [Schurz 1998a](#) , [2024b](#)). Representatives of the empirical-analytical program are (unfortunately) the only ones striving for a clear division between description and value judgments. In hermeneutic approaches, by contrast, descriptive statements and value judgments are often deliberately not distinguished from each other (cf. [Hilgendorf 2000](#) , 39ff; [Dupré 2007](#)), and in critical-dialectical approaches, certain emancipatory norms are presented in a quasi-scientific make-up.

Our analysis has straightforward consequences on some debates on the role of values in science ([Kincaid et al. 2007](#) ; [Carrier et al. 2008](#)). A significant number of contributors in this debate have criticized the ideal of value-free science. Their objections can be resolved by our analysis of means-end inference and the resulting requirement of separating empirically confirmed knowledge from value assumptions. Dupré ([2007](#) , 30–35), for example, argues that science needs to employ value-loaded terms in order to be practically relevant. Dupré is right insofar as the derived norms recommended by scientists employ value-loaded terms. But he seems to overlook that the scientist (qua scientist) recommends these derived norms only *hypothetically* and instrumentally, as means for certain fundamental norms that are given by society.

As emphasized in [prop. 2.5-1](#) , the requirement of value neutrality applies only to the scientific *context of justification* (CJ): it must be independent of fundamental, extra-scientific value assumptions. By contrast, the *context of origination* (CO) inevitably involves extra-scientific values that influence the selection of research questions. However, proponents of scientific value-involvement – such as Douglas ([2000](#)) – argue that even the CJ is influenced by extra-scientific values. In support of this position, the following argument is often made: First, the formulation of the research question in the CO is guided by non-epistemic interests. Second, in the CJ, the research question significantly shapes the choice of scientific method. This shows that non-epistemic interests can have an *indirect* influence on methodological choices, or more broadly, that choices within the CO can have an impact on the CJ.

The argument of indirect influence is certainly correct, but may create confusion. Once a research question is precisely specified, the choice of the optimal method for finding an insightful answer becomes a *purely epistemic* matter. As shown in Schurz ([2024b](#) , 164–167), this also applies to the determination of the “significance level” that is often-cited in this context. In Parker's ([2024](#)) words, the specification of the research question projects an initially non-epistemic interest into a purely epistemic research task. We illustrate this with an *example* : A research project on effective treatments for depression may be guided by non-epistemic interests of the pharmaceutical industry. As a result, the research question may be limited to pharmaceutical interventions: Which antidepressant drugs produce the best outcomes (operationalized in terms of efficiency, sustainability, etc.)? For a research question that has been specified in this precision, the choice of the optimal method (data collection, statistical procedures, etc.) no longer depends on non-epistemic interests. This holds true even if it turns out that, for certain types of depression, psychological therapies are more effective than antidepressant medication. In that case, a properly conducted study would simply yield a largely negative result – that is, antidepressants would be shown to have only limited effectiveness. ⁵

The example of pharmaceutical depression studies leads to another important insight. In our example, it becomes evident in the CJ that restricting the investigation of depression therapies to pharmacological treatments impedes research, because it neglects important causal factors that plausibly lie within the psychological domain. In such a case, the selection of research parameters initially made in the CO must remain open to later revision based on findings in the CJ, even if such revisions run counter to the extra-scientific goals of the research project. Otherwise, one can no longer speak of the CJ being independent from external value commitments.

As a final application of our concept of value neutrality, we turn to the so-called *new demarcation problem* ([Wilholt 2009](#) ; [Holman/Wilholt 2022](#)). This challenge arises for all opponents of value neutrality. The challenge consists in the fact that, on the one hand, advocates of value-laden science massively criticize the capitalist value bias in industry-sponsored studies, while at the same time maintaining that value influences on scientific knowledge are both inevitable and desirable. But this implies that they should regard not only their own (typically left-leaning) values as legitimate, but also those underlying industry-funded

capitalist research (cf. also [Henschen 2021](#) , 20). Most proponents of value-laden science seek to avoid this implication – but how? Some authors have proposed resolving the new demarcation problem by suggesting that scientists should rely on a specific set of preferred values – namely, those reflecting democratic majoritarian values forming the normative foundation of our society. ⁶

Le Bihan ([2024](#)) has raised a fundamental objection to the thesis of democratic value-dependence: modern democratic societies are polarized, with left- and right-leaning segments of the population holding opposing value preferences on many issues – from immigration to climate policy. In this context, the proposal to ground science in “democratic values” offers little guidance, as it leaves open on which side of the political spectrum these “democratic values” are supposed to lie. Moreover, the proposal carries a disturbing implication: in times of political upheaval – such as the current situation in the United States – major scientific findings and textbooks would have to be rewritten, since they would depend on democratic majority values that have now fundamentally shifted. Fortunately, the prevailing conception of science in modern democratic societies has largely freed itself from such political entanglements – and it should remain that way.

From a value-neutral perspective, the question of the “right values” is not a fundamental problem, even in times of value polarization. This is because the goals to which scientists relate their recommendations for means are determined by society. If society is in a state of *value polarization* , then value-neutral policy advice requires conducting hypothetical cost-benefit analyses for *multiple, diverging* value preferences that reflect the political positions of relevant segments of the population (cf. [Carrier 2022](#)). This was also a key element in the learning process reflected in the recommendations of the German Leopoldina commissions during the COVID-19 pandemic (see [Wiesing et al. 2021](#) ; [Schurz 2024b](#) , 158f).

2.10.2 The Argument of Inductive Risk

A much-debated objection against scientific value neutrality is the argument from inductive risk (AIR), originally proposed by Rudner ([1953](#)) and further developed by Douglas ([2000](#)). The AIR assumes that scientific hypotheses are typically supported by empirical evidence only with a certain degree of probability. When scientists accept a hypothesis H as true, they thereby incur a certain “inductive risk” – the risk of error. How small must H’s risk of error be in order for it to be reasonable to accept H as true – that is, to rely on H and be willing to act on the basis of H? The *risk* associated with accepting a hypothesis H depends on two factors: (i) the *error probability* of H (i.e. the probability of $\neg H$ given the total evidence) and (ii) the practical *cost* incurred by accepting H when H is false. While (i) is an epistemic, (ii) is a non-epistemic question, depending on extra-scientific values. H’s total risk of error is given as the product of (i) and (ii). It follows that for hypotheses with practical relevance, their acceptability as a basis for practical action depends inevitably on extra-scientific values.

Quite a few authors see the AIR as convincing evidence for the influence of non-epistemic values on the CJ ([Holman/Wilholt 2022](#) , [Brown 2024](#)). In reality, however, the AIR is weaker than it sounds – a “storm in a teacup.” The AIR can be defused by the demand of *uncertainty transparency* in science, that is, by explicitly stating the error probabilities of scientific claims. This demand not only serves the goal of value neutrality but also helps prevent problematic consequences of expertocracy. We illustrate this with the example of the prediction of *earthquake safety* . Let H be the hypothesis that (in a given region over the next few days) no earthquake of magnitude greater than 6.5 on the Richter scale will occur. Even if the error probability of this hypothesis is only 5%, many affected people would be poorly advised to act on it – at least those whose lives would be seriously endangered in case of error, because their houses are poorly built and at risk of collapse at this level of seismic activity. It is therefore undoubtedly correct that the question of whether to accept non-H rather than H – and thus whether or not to evacuate – depends on non-epistemic cost-benefit considerations (for detailed calculation s. [Schurz 2024b](#) , 160–167 and [sec. 5.9.5](#)).

But does this really mean that scientists should make such value judgments on behalf of the affected people, as proponents of the AIR would have us believe? Certainly not. Instead, the error risks involved should be explicitly stated and explained. The L’Aquila earthquake illustrated this in a tragic way: In October 2009, six earthquake experts and a government official in Italy were convicted in the first instance and sentenced to several years in prison because they had failed to predict an earthquake that occurred on April 6, 2009, in L’Aquila, in which more than 300 people died – instead, they had given the all-clear ([Cocco et al.](#)

[2015](#)). The seismologists defended themselves against the verdict by pointing out that they had not issued a categorical all-clear, but rather a more cautious probabilistic judgment, namely that the observed slight increase in seismic activity was *not significant* , with an estimated error probability of 5%. As a result, the verdict was overturned in 2014; only the government official who had communicated the “all-clear” message to the media was ultimately convicted.

The AIR is based on the mistaken belief that it is the task of science to provide the public with categorical knowledge claims. Quite the contrary: objectively value-neutral science requires that for practical uncertain knowledge claims, the estimated error probabilities should always be reported – an insight first emphasized by Jeffrey ([1956](#) , 237). Referring to Schurz ([2013](#)), both Betz ([2017](#)) and Henschen ([2021](#)) have likewise proposed to solve the problem of inductive risk by making error probabilities explicit. Douglas and like-minded authors attempt to refute this position by arguing that people expect statements from the “authority of science” with clear-cut practical implications – such as giving an “all-clear” ([Douglas 2000](#) , 563; [2009](#) , 135; [Wilholt 2009](#) , 94). But scientists must not give in to such false expectations; rather, it is part of their responsibility to enlighten the public about the limits of the knowable and the risk of error.

However, the critics are right to point out that a statistical significance statement, hardly understandable for laypeople, is not sufficient for such public education. We suggest that error probabilities should be made comprehensible to non-experts by means of *hypothetical cost-benefit considerations* . In the earthquake example, the 5% error probability could be explained as follows: *If* the negative consequences of an earthquake occurring while one remains in their home are estimated to be more than 20 times greater than the costs of a temporary evacuation, *then* it would be advisable to evacuate.

2.10.3 Meta-ethical Justification of Value Neutrality and the Difference between Experience and Valuation

The value-neutrality requirement is based on the thesis that fundamental norm and value sentences are not capable of being justified by empirical science. In these subsections, I discuss this thesis in the context of *meta-ethics* , which is that part of ethics that investigates the meaning and justifiability of ethical statements. First and foremost, value neutrality requires the separation of descriptive statements from value and normative judgments. In formal scientific languages, this separation is already ensured by the differing forms of expressions. However, in everyday language, a problem arises: many terms in our everyday vocabulary carry *both* factual and evaluative content, often closely intertwined, making conceptual separation difficult. Some critics of value neutrality, such as Putnam ([2002](#) , ch. 2) and Dupré ([2007](#)), have claimed that in the case of so-called thick concepts such as “cruelty” or “rape,” such separation is impossible. Yet, as Hare ([1981](#) , 74f) has already argued, this separation is more a *matter of will* than of ability, and with the help of suitable linguistic tools, it is always possible.

The standard meta-ethical justification for the value-neutrality requirement is the *fact – value* or *is – ought* dichotomy , which goes back to David Hume and asserts that value or normative statements cannot be logically or analytically derived from purely descriptive statements ([Hume 1739](#) /1740, 177f). Attempts to prove this prima facie highly plausible thesis by means of modern logic were confronted with unexpected problems, such as the paradox of Prior ([1960](#)) (cf. [Stegmüller 1979a](#) , 183–186). As a result, the dichotomy thesis was called into question by a number of philosophers ([MacIntyre 1981](#) ; [Putnam 2002](#)) and could only be placed on a solid foundation through more recent logical investigations. ⁷ Schurz ([1997](#)) proves that it is only possible to derive nontrivial ethical statements from descriptive premises if one already presupposes so-called *is-ought bridge principles* in one’s ethical theory, such as the utilitarian principle “Whatever is desired by the majority of humans, is ethically good, or ought to be realized.” Is-ought bridge principles are not analytically valid, but are synthetic assumptions belonging to one’s accepted ethical system and are typically controversial. Therefore, norms or values cannot be justified by descriptive (non-ethical) premises without already assuming hidden ethical principles.

There exists a variety of different meta-ethical positions (e.g. utilitarianism and a priorism), but none of them stands in a substantial conflict with Hume’s Is-ought thesis. ⁸ The only exception is the position of so-called *value empiricism* (cf. [Czaniera 2001](#) , ch.1–2). This position assumes a special kind of “value experience,” which, although distinct from biological sensory experience, is analogous to it in that it relates to individual experiences and can serve as a basis for testing in the value sciences. This position has been

advocated by, among others, Firth ([1952](#)), Weingartner ([1978](#) , §7.15 ff), and more recently by Anderson ([2004](#)). In the following, I will use two arguments to show that such “value experiences” behave quite differently from genuine sensory experiences: the normal conditions argument and the independence argument.

A The *normal conditions argument* says that intersubjectively deviating perceptual judgments can almost always be traced back to *defects* in the normal conditions of the observation. But this does almost never apply to intersubjectively deviating value judgments. Let us consider, for example, the presumably true observation sentence: “There is a fire there.” The thesis of the *intersubjectivity* of such an observation sentence implies that any or almost any person would agree with this sentence, provided the *normal conditions* of observation explained in [sec. 2.9.2](#) are fulfilled. If a person within visual range of the fire would really repeatedly and persistently claim “but I see no fire,” then we would have every reason to assume that there is a defect in the person’s normal conditions, and we would be able to find out this defect by independent tests. Such a person would most probably be blind or brain-damaged, or could be in a state of extreme hallucination. For this reason, we would also be entitled to try to *cure* such a person of their observational defect.

But now let us assume that an adult person, under the same normal conditions of observation, systematically *disagrees* with a basic value sentence such as “It is good that *this* person is being kept from dying by being fed,” although we assume that 99% of all persons do agree with it. Then we would *not be entitled* to assume that there is a defect in the ‘normal moral condition’ of this person, that is, that this person is morally inferior and should be cured, just like the blind should be cured of their blindness. Rather, we must *respect* this person’s position *prima facie* as a deviating value position and also – as far as it does not infringe the rights of others – accept it. It is normal in Inuit society, for example, not to give food to a person who is so weak with age that they cannot feed themselves, but to let them die instead, because this is seen as a natural sign that their spirit wants to leave this world. Should we therefore declare the morals of the Inuit to be ‘morally defective,’ and try to ‘cure’ them of their defects? If we treated basic value sentences exactly like observation sentences, we could find ourselves moving dangerously close to medieval religious moral practices.

I do not want to deny with this thesis that there are indeed cases of *psychopathological* blindness of conscience, which are in need of a psychological cure – but these cases are not the *rule*. An extreme claim such as “I love killing other people” would rightly cause anybody to react by thinking “this person has a defect” – but this situation applies only to *very few* deviating basic value sentences, whereby the analogous situation in the case of deviating judgments of perception is the *rule*. Quite in this sense, Prim and Tilman ([1979](#) , 93f) report an experiment on the objectivity of the observation sentence: “there is a piece of chalk in front of me,” which they conducted on their students: the students reacted very indignantly, and declared that dissent in such questions was a problem for an *optician*, and not one of philosophy.

B The *independence argument* goes back to Toulmin ([1950](#) , 127) and runs as follows: Our ‘value experiences’ are *dependent* on our *moral background assumptions*. Our perceptual impressions, on the other hand, remain the same, even when our descriptive background assumptions change radically. For example, we see the Sun make its way across the sky day in, day out. Let us assume that up to now we have held the geocentric view (the Sun revolves around the stationary Earth); but now, after detailed study of Kepler and Newton, we have changed to the heliocentric view (the Earth revolves around the stationary Sun). There has now been a radical change in our background theory – but as already explained in [sec. 2.9.1](#) (contra-3), our visual experience is changed *in no way* by this fact: just as before, we see the Sun drawing an arch across the sky every day. Our visual experience is independent of acquired background assumptions.

This is quite different in the case of value-based moral sentiments. If I see somebody donating a sum of money to the poor, I may spontaneously have the moral sentiment: *that was a good deed*. If I find out, however, that the donor is a politician about to stand for election, who has invested this donation as part of a calculated plan to win votes, but is otherwise more concerned about the well-being of the rich than of the poor, then a quite different moral sentiment will arise within me: I now see the politician as a *hypocrite*.

The normal conditions and the independence arguments strengthen our thesis that an empirical justification of value or norm sentences in analogy to empirical science is not possible. Value and norm-

based judgments are not based on sensory experiences, but on subjective interpretations of experiences that depend on cultural background values and emotional impressions arising in the context of the perceived situation.

2.11 The Justification of Induction: An Unsolvable Problem?

2.11.1 Is Induction Rationally Unjustifiable?

As mentioned in [sec. 2.7.1](#), Hume's problem of induction consists in the question of how inductive reasoning can be rationally justified. David Hume made it clear to the philosophers of his time that, *first*, inductive inferences cannot be justified by means of *logic*, because it is logically possible that the future (or the unobserved) is radically different from the past (or from the observed). However, Hume went much further: he cast doubt on the rational justifiability of induction as a whole (1748, sec. 4). Induction, *second*, also cannot be justified by observation, because in inductive inferences one infers that which has *not yet* been observed. And *third*, inductive inference also cannot be justified by induction, as this would lead us straight into a *circular argument*: if we think inductive inference is reasonable because it has been corroborated so far, then we are assuming that what has been corroborated so far will – by inductive inference – also be corroborated in the future. In such a circular argument, we assume exactly the thing we want to show, and so we show nothing at all. All three traditional methods of scientific justification – logic, experience, and induction – thus appear to fail when it comes to justifying induction itself. Hume was also the first to show that, *fourth*, a reformulation of the induction problem using probability theory does not help (1748, sect. 6). If we argue that inductive inferences do not always, but at least *with great probability* lead us from true premises to true conclusions, then we are assuming a probabilistic induction principle, according to which the *event frequencies* observed so far may be inductively transferred to the future or to non-observed cases.

These were the reasons which led Hume to his deeply *skeptical* thesis, that inductive inference is not capable of being rationally justified, but is based merely on psychological *habit* (1748, sect. 5). Finding a satisfactory solution to Hume's induction problem would be of the greatest importance. And yet, to this day, there is no generally accepted solution to the problem. ⁹We briefly discuss five solution attempts that have been suggested in the last few decades but could not find general acceptance:

1 *Induction is not needed in science?* Popper's attempt to circumvent the induction program was to argue that scientific methodology doesn't need induction. In [sec. 2.7.2](#), we have seen, however, that Popper's argument fails, because Popper's own methodology of "corroboration" contains in its heart the principle of epistemic induction.

2 *Induction as rational by definition?* Proponents of the analytic account (e.g. [Strawson 1952](#), 257; [Ayer 1956](#), 74f, [Pollock 1974](#), 204) have argued that induction is rational by the very definition or meaning of rationality. Recall, however, that by a rational justification we don't mean a mere 'intuition of rationality' (which is then declared as being 'true by definition'), but a demonstration that inductive inferences are truth-conducive, in the sense of optimizing our success in delivering true predictions. This cannot be shown by a mere 'definition of rationality,' no more than diseases can be cured by a 'definition of health.' Nor is Strawson's claim (ibid., 249) correct that we cannot imagine a situation in which inductive inferences fail (cf. [Salmon 1974](#)). We can not only conceive of such situations in thought (consider *Alice in Wonderland*); in fact, millions of people *do* believe in the superiority and even in the 'higher rationality' of non-inductive 'methods,' be it clairvoyance, fortune-telling, or God-guided inner intuition. So, a satisfactory justification of induction would not only be of high philosophical, but also of high *cultural* importance, as part of the enterprise of explaining scientific rationality.

3 *Rule-circular 'justification' of induction by induction?* The circularity involved in the 'inductive justification of induction' has been called *rule-circularity* ([Ladyman/Ross 2007](#), 75), because here, the conclusion (the reliability of induction) is not an element of the premise set, but is presupposed by the underlying inference rule (namely induction). Several philosophers expressed the hope that rule-circular arguments have epistemic value and might offer a solution to the problem of induction. ¹⁰The following counterexample by Salmon ([1957](#), 46) destroys this hope. It shows that the same type of rule-circular

argument that ‘justifies’ the reliability of induction can be used to ‘justify’ the rule of counterinduction, which says, roughly speaking, that one should predict the opposite of what one has observed:

(2.11-1)

Rule-circular justification of induction :

Premise : Past inductions have been successful.

Therefore, by rule of induction:

Inductions will be successful in the future

Rule-circular justification of counterinduction :

Premise : Past counterinductions have not been successful.

Therefore, by rule of counterinduction:

Counterinductions will be successful the future.

The two circular arguments share exactly the same structure, the premises of both arguments are true, and yet they lead to opposite conclusions. This proves that such rule-circular arguments are epistemically worthless pseudo-justifications. A generalization of this refutation for inferences to the best explanation is found in Douven ([2011](#)), and for arbitrary rule-circular arguments in Schurz ([2024a](#) , [sec. 4.3](#)). Rule-circular justifications can also be given for dangerous fundamentalistic rules, such as the “rule of blind trust in my leader,” abbreviated (BTL), whose rule-circular justification goes as follows: “I believe in the reliability of the rule BTL because this was assured to me by my leader.”

4 Induction justified by uniformity of nature? John Stuart Mill ([1865](#) , III .3.1) and Bertrand Russell ([1912](#)) held the view that the reliability of induction could be justified by assuming the uniformity of nature (i.e. nature’s future resembles its past). The main problem with this approach is that it also leads to a *justificatory circle* , since the assumption of uniformity can itself only be justified by an inductive inference. Note that as in ([2.11-1](#)) above one may give a corresponding circular justification of counterinduction by counteruniformity: counterinduction is justified by counteruniformity which is in turn justified by counterinduction from past observations. A “localized” version of the uniformity-account has been proposed by Norton ([2003](#)). For Norton, inductive reasoning is not governed by formal but by domain-specific or “local” inference rules such as “if some samples of Bismuth melt at 271 ° C, then all samples of Bismuth will do so.” Local inductions are in turn justified by local uniformity assumptions, for example, “samples of the same element agree in their physical properties.” However, Norton’s proposal is subject to exactly the same charge of circularity as the Mill-Russell approach, since local uniformity assumptions must be justified by inductive inferences (cf. [Kelly 2010](#) ; [Schurz/Thorn 2020](#)).

5 Induction justified by inference to the best explanation? Several authors ([Harman 1965](#) ; [Armstrong 1983](#) , sec. 6; [Lipton 1991](#) , 69) have argued that inductive inferences are justified because they are instances of inferences to the best explanation (IBE). Recall from [sec. 2.7.3](#) that an IBE infers from an observed phenomenon that the premises of its best (available) explanation are probably true. An inductive generalization – so the argument goes – is an instance of an IBE , because the best explanation of a regularity observed so far (e.g. “so far all observed ravens were black”) is the corresponding lawlike generalization (“It is a law that all ravens are black”).

The major problem of the IBE-account is, however, that the justification of IBEs is even more difficult than the justification of induction. This is so because IBEs involve even more uncertainties than inductive inferences (recall [sec. 2.7.3](#)). Analogous justification strategies to those discussed above for induction have been proposed for IBEs (cf. [Bird 1998](#) , 171). Armstrong ([1983](#) , 53), for example, has argued that IBE is rational ‘by definition.’ Lipton ([1991](#) , 167ff) and Psillos ([1999](#) , 82) suggested to justify IBE in a rule-circular way. But we have seen that both strategies are bound to failure. Moreover, it is unclear why one should regard the explanation of “All so-far observed Fs have been Gs” by “Because it’s a law that all Fs are Gs” as a *best* explanation, if one does not accept the inductive premise that the world is uniform. If one assumes instead that from time to time the laws of nature undergo radical changes, the explanation “Because it was *so-far* a law that Fs are Gs” seems to be the better explanation.

2.11.2 The Optimality of Meta-induction

Due to the many failed attempts to solve the problem of induction, quite a few contemporary philosophers consider it unsolvable. ¹¹ We do not share this view. We agree with Hume's skeptical diagnosis that no noncircular argument can establish that induction is *reliable*, that is, will be predictively highly successful, whatever the future looks like. However, there is an alternative justification strategy that is not blocked by Hume's argument: namely, the attempt to show that induction is an epistemically *optimal* prediction method, in the sense of being most successful among all prediction methods that are *accessible* to a given agent. Even in radically skeptical scenarios in which all methods of prediction are unreliable, induction can still be optimal in the sense of 'being the best of a bad lot.'

A noncircular optimality justification of induction was developed by Schurz (see, e.g., [Schurz 2008b](#), [2019a](#)). A precursor to this approach is Hans Reichenbach's justification of induction as the "best alternative" ([Reichenbach 1949](#), §91). However, Reichenbach failed to develop a robust optimality argument in favor of induction, as he based his argument on *object-induction*, that is, induction aimed at predicting events. Yet, nothing can rule out the logical possibility of a clairvoyant who predicts a random sequence with a significantly higher success rate than the scientific inductivist (see [Skyrms 1999](#), sec. III.4). By contrast, the approach developed by Schurz is based on the idea of *meta-induction*.

The meta-inductivist applies the inductive method to all accessible prediction methods and attempts to derive an optimal prediction strategy based on their observed track records. The simplest meta-inductive method is *Imitate-the-best*, abbreviated as ITB. It predicts what the actually best non-MI method predicts (resolving ties according to a given ordering of accessible methods). ITB's success will be optimal in the long run if, after some finite time, there is a *unique* best method that ITB imitates forever (cf. [Schurz 2019a](#), [sec. 6.1](#)). Interestingly, ITB is *not* universally optimal. ITB's optimality breaks down when it plays against 'deceiving' methods that lower their success rate as soon as their predictions are imitated ([Schurz, 2008b](#), sec. 4). A realistic example is predictions of stock values in a 'bubble economy:' here, the prediction that a given stock will yield a high return leads many investors to buy that stock, causing it to crash by doing so (since the stock lacks sufficient economic support).

A more refined meta-inductive strategy is attractivity-weighted meta-induction. This meta-inductive method, abbreviated as AW, predicts a weighted average of the predictions of all methods accessible to it, with weights determined by the methods' past success rates (overview in [Schurz 2019a](#), sec. 6–7). Building on results from mathematical learning theory ([Cesa-Bianchi and Lugosi 2006](#)), Schurz ([2008b](#)) could show that AW is indeed universally optimal: in all prediction games with arbitrary sequences of events and a finite number of available methods, AW's predictive long-run success is greater than or equal to the maximal long-run success of all accessible methods ([Schurz 2008b](#), theorem 4; [2019a](#), theorems 6.8–6.11). Even in 'radically nonuniform' scenarios in which the success rates of the leading players permanently oscillate around each other, AW's success rate converges to the maximal success rate (or to a greater value), with a small short-run loss that converges to zero with an increasing number of predictions.

The precise formulation of the meta-inductive optimality theorem runs as follows. Let $(e) =_{\text{def}} (e_1, e_2, \dots)$ be a sequence of events e_n ordered along a discrete time $n = 1, 2, \dots$, whose possible values are coded by real numbers in the interval $[0,1]$. Let $M = \{M_1, \dots, M_m, AW\}$ be a finite set of (m) simultaneously accessible prediction methods together with AW. The methods $M_1, \dots, M_m$ may be given as computational algorithms, real-life experts or even (purported or real) clairvoyants; their "simultaneous accessibility" means that the epistemic agent can track and compare their success records, either by observation or by simulation. The pair $((e), M)$ is called a *prediction game*. In each round n of the game, it is the task of each method to deliver a prediction $\text{pred}_{n+1} \in [0,1]$ of the next event e_{n+1} of the sequence. The deviation of the prediction pred_n from the event e_n is measured by a normalized loss function, $\text{loss}(\text{pred}_n, e_n) \in [0,1]$. The optimality theorem below holds for all *convex* loss functions (a generalization is possible that even holds for all loss functions). The predictive success of a method M at a given time point n is evaluated by its *success rate* $\text{suc}_n(M)$, defined as the sum of M 's scores achieved until time n divided by n (where M 's score achieved at time n is $1 - \text{loss}(\text{pred}_n, e_n)$). The attractivity of a method M for AW at time n , $\text{at}_n(M)$, is defined as $\text{suc}_n(M) - \text{suc}_n(AW)$. The prediction of the meta-inductive method AW for the next time $n+1$, $\text{pred}_{n+1}(AW)$, is the normalized weighted average of the predictions of the methods $M_1, \dots, M_m$ for this time with weights

defined as $w_n(M_i) =_{\text{def}} e^{\eta \cdot n \cdot \text{at}_n(M_i)}$, where $\eta = \sqrt{8 \cdot \ln(m)/(n+1)}$. “ maxsuc_n ” denotes the maximal success rate of the accessible methods at time n . With these definitions, the optimality theorem runs as follows:

(Prop. 2.11-1) *Universal optimality of AW*: ¹²

For every prediction game $((e), \{M_1, \dots, M_m, AW\})$:

- i. (Short run) $(\forall n \geq 1:) \text{suc}_n(AW) \geq \text{maxsuc}_n - 1.77 \cdot \sqrt{\ln(m)/n}$.
- ii. (Long run) $\lim_{n \rightarrow \infty} (\text{suc}_n(AW) - \text{maxsuc}_n) \geq 0$.

In the short run, weighted meta-induction may suffer from a possible loss (or “regret”) compared to the leading method, caused by the fact that AW must base her prediction for the next time on the *past* success rates of the accessible methods. However, proposition 2.11-1(i) states a worst-case upper bound for this loss, which is small if the number of rounds n is not smaller than the number of candidate methods m , and which quickly converges to zero when n grows large.

AW cannot get deceived by a method because it never favors only one method but a weighted average of all attractive methods. A second important feature of AW is that it tends to *ignore* methods that are constantly worse than itself, since exponential weights with negative attractivities quickly converge to zero. This ignorance-tendency is necessary for AW’s long-run optimality and a crucial distinction of AW compared to success-weighted meta-induction that uses simple success rates as weights and is provably nonoptimal (Schurz 2019a, sec. 6.8.1). It follows that in games in which one method dominates its competitors, the weight of the best method will quickly converge to 1, and AW’s behavior will converge to that of ITB.

The optimality result holds independently of any assumptions about the possible worlds or their prior probabilities. This means that meta-induction possesses a mathematical a priori justification. But what does this result imply for the justification of object-induction? In possible worlds that contain genuine clairvoyants, object-induction cannot be optimal; in such worlds, meta-induction does not favor the object-inductivists, but rather the clairvoyants. However, the a priori justification of meta-induction provides us with the following noncircular *a posteriori justification of object-induction*: we know from experience that object-inductive prediction methods have so far outperformed all other available prediction methods. Therefore, we are justified – by meta-induction – in continuing to use object-induction in the future. This justification is, unlike the “inductive justification of induction,” no longer circular, because we have an independent a priori justification of meta-induction.

In the meantime, a lively debate has emerged regarding the prospects of the meta-inductive optimality justification, in which further challenges have been raised and solutions proposed. ¹³In Schurz (2024a), optimality justifications are extended to the entire field of epistemology.

2.12 Introductory Literature and Exercises

2.12.1 Introductory Literature

On the method of rational reconstruction, see Stegmüller (1969 ff, vol. IV.1). A classic on the German value judgment debate is Albert and Topitsch (1971, ed.), for more recent debates, see Kincaid et al. (2007) and Carrier et al. (2008); a recent introduction is Elliott (2022). A collection on the question of the unity of sciences is Rahman et al. (eds., 2004). Psillos (1999) offers an excellent introduction to the empiricism versus realism debate on philosophy of science. Recent overviews on the metaphysics of science are Schrenk (2017) and Hüttemann (2021). Rock (1984) introduces into the cognitive psychology of perception. Swinburne (1974) and Henderson (2018a) offer introductions to the problems of induction. Lipton (1991) is a classic on inference to the best explanation and Niiniluoto (2018) is a good introduction to abduction in science.

2.12.2 Questions and Exercises

Sec. 2.1

1. What is meant with “context of discovery” versus “context of justification?”
2. Explain the normativist argumentation and its criticism.
3. What is the supreme epistemic goal of science? Does it presuppose metaphysical realism?
4. Explain the method of rational reconstruction.

Sec. 2.2–2.4

1. Explain the five common epistemological assumptions of the sciences.
2. In which sense are minimal realism and minimal empiricism “minimal?”
3. What is the difference between falsifiability and fallibility?
4. Explain the common methodological features of sciences.
5. To what extent are the concepts of “natural sciences,” “human and social sciences,” and “humanities” ambiguous?
6. How do empirical, experimental, and dissecting sciences differ from one another in regard to the degree of their empirical invasiveness? Classify educational science, human medicine, and cultural studies within this classification.

Sec. 2.5–2.6

1. Explain the means-end inference using examples from your area of study.
2. Explain the suggested formulation of the value-neutrality thesis and all concepts contained in it.
3. Discuss: How is the requirement of “value neutrality” related to that of “political correctness?” Are the two in conflict?
4. Discuss physical reductionism and its criticism. What is “supervenience?”

Sec. 2.7

1. Which kind of inference (deductive, inductive, or abductive) is being used in the following examples?
(i) Most light switches are working. I am pressing the switch. Therefore, the light will go on. (ii) All humans are mortal. Aristotle is human. Therefore, Aristotle too is mortal. (iii) My refrigerator has always worked well up to now. Therefore, I will be able to rely on it tomorrow. (iv) Every time Gregor talks about his brother, his facial expression becomes strained. Therefore, he has a problem with his brother.
2. Explain the principles of methodological, logical, and epistemic induction. What are the Popperian arguments against the first two kinds of induction?
3. Give an example of an inference to the best explanation.
4. Which inferences in 1 are monotonic and which are non-monotonic?

Sec. 2.8–2.9

1. What is the difference between metaphysical and hypothetical-constructive realism?
2. Explain the argument of radical constructivism against realism and its criticism.
3. One person could fail to notice another because (a) he is deliberately ignoring her, (b) he is blind, (c) he is asleep, or (d) it is pitch dark. Which normal condition of observation is infringed in each case?

Sec. 2.10–2.11

1. Explain Habermas’ threefold division of scientific disciplines according to their respective interests. How is this view criticized?
2. Discuss the normal conditions and the independence argument.
3. Recapitulate David Hume’s four arguments against the possibility of justifying induction.

4. Explain the structural equality of the inductive justification of induction and the counterinductive justification of counterinduction.
5. What is a meta-inductive prediction method? In which sense is attractivity-weighted meta-induction optimal?

Notes

1. Cf. Oppenheim/Putnam ([1958](#)), Causey ([1977](#)), Hofer/Smeenk ([2016](#)); Hüttemann ([2021](#) , ch. 5–6). [↩](#)
2. Cf. Humphreys ([1997](#) , 10), Ladyman/Ross ([2007](#) , sec. 1.6), and Hüttemann ([2021](#) , 151–155). [↩](#)
3. For arguments in this direction cf. Chalmers ([1986](#) , [sec. 3.4](#)), van Fraassen ([1980](#) , 56ff), Albert ([1985](#) , sec. II.8), Brewer/Lambert ([2001](#) , S182f). [↩](#)
4. For example, Putnam ([1962](#)), Shapere ([1982](#)), Kuipers ([2000](#)), Giere ([1999](#)), Adam ([2004](#)). [↩](#)
5. It is well known that the so-called *positivity bias* in scientific research often leads to the non-publication of negative results, which causes significant distortions in the research record ([Ioannidis 2005](#) ; [Schooler 2011](#)). One example is evidence of bias in pharmaceutical studies in favor of their industrial sponsors ([Brown 2008](#)). However, such distortions are not an inherent feature of science but rather a *defect* of scientific practice under external pressures. In response, a number of scientific journals have introduced measures to ensure the publication of negative research findings. [↩](#)
6. For example, Elliott ([2017](#)), Schroeder ([2021](#)), Alexandrova/Fabian ([2022](#)). [↩](#)
7. S. Pigden ([1989](#)), Schurz ([1997a](#)), Russell ([2022](#)); overviews in Hudson ([1969](#)) and Pigden ([2010](#)). [↩](#)
8. Overview of meta-ethical positions is, for example, found in Frankena ([1963](#)), Kutschera ([1982](#) , sec. 2), Schurz ([1997a](#) , sec. 11.3), and Sayre-McCord ([2023](#)). [↩](#)
9. Cf. Stegmüller ([1977](#) , vol.II, ch. 4), Swinburne ([1974](#)), Rescher ([1980](#)), Howson ([2000](#)), Ladyman ([2002](#) , ch. 1), Schurz ([2019a](#) , ch. 2–4). [↩](#)
10. For example, Braithwaite ([1974](#)), Black ([1974](#)), van Cleve ([1984](#)), Papineau ([1993](#) , sec. 5), Goldman ([1999](#) , 85), and Psillos ([1999](#) , 82). [↩](#)
11. See, e.g. Stegmüller ([1977](#) , vol. II, ch. 4), Greco ([2005](#) , 265); Williamson ([2023](#) , 451). [↩](#)
12. The loss term in (i) simplifies the loss term in th. 6.9(ii) of Schurz ([2019a](#)), since $\sqrt{2} + 1/\sqrt{8} < 1.77$. [↩](#)
13. See, for example, Arnold ([2010](#)), Henderson ([2018a](#) , sec. 7.3), Sterkenburg ([2019](#)), Henderson ([2023](#)), Douven ([2023](#)), Briesen ([2026](#)), Shogenji ([2023](#)), Wolpert ([2023](#)), Pitts ([2023](#)), Kimling ([2023](#)); Replies in Schurz ([2021b](#) , [2023a](#) , 2023b, 2024c, 2025). [↩](#)

3 The Conceptual Toolkit

Language, Logic, and Probability

DOI: [10.4324/9781003675075-3](https://doi.org/10.4324/9781003675075-3)

This chapter has been made available under a CC BY-SA license.

3.1 Kinds and Classification of Concepts

The basic logical building blocks of any language are its concepts. A concept is understood here as a linguistic expression *together* with its meaning. Grammatically, concepts are *beneath* the level of the sentence – they are *subsential*. “Grass” and “green” are concepts; “grass is green” is a sentence. Though concepts signify something, they cannot in themselves be true or false; only sentences composed of concepts can. Concepts may be *atomic* or *complex*; for example, “grass” is an atomic, while “green grass” is complex. Sentences are composed of concepts, and the nature of a sentence depends on the nature of its concepts. So, we will *first* deal with the various different kinds of scientific concepts, before dealing with the various different kinds of scientific sentences. Concepts may be classified in three respects: first according to their logical type, second according to their content type, and third according to their gradation (or scale) type.

3.1.1 The Logical Type of Concepts

Logical grammar classifies concepts according to their logical-semantic function. [Fig. 3.1-1](#) presents an overview.

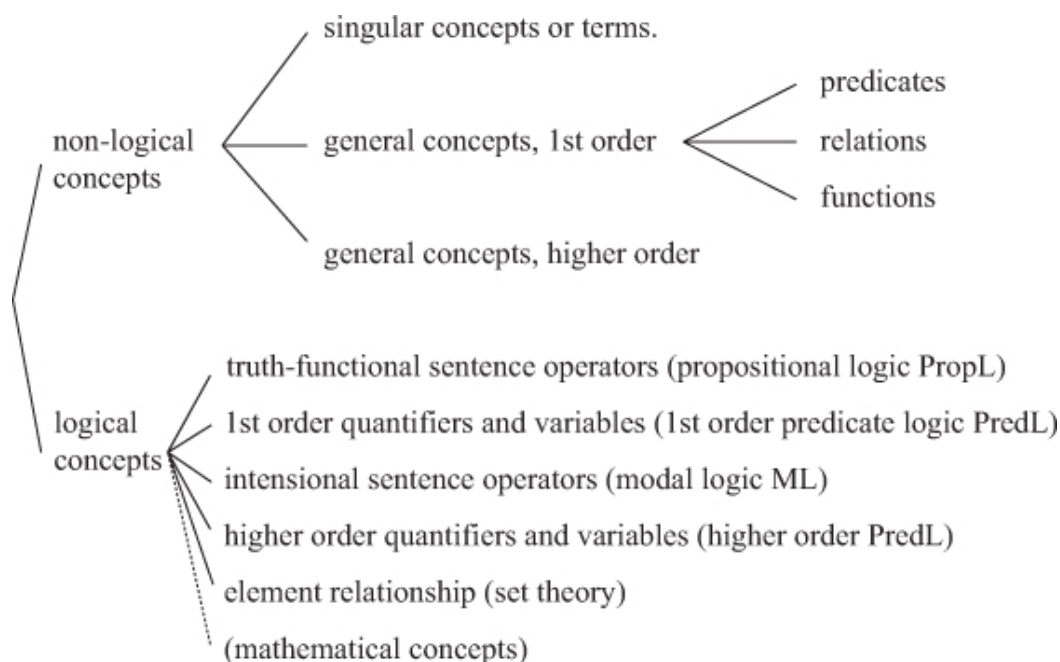


Figure 3.1-1 Division of concepts according to their logical type. [↩](#)

1. *Nonlogical concepts* denote or signify something in the world . By contrast, logical concepts have a purely *structural* function.

1.1. *Singular concepts* or *terms* always designate a single *individual* (or spatiotemporal location). Singular terms subsume (a) *proper names* (e.g. “Saul Kripke,” “Salzburg”), (b) *ostensive* or *indexical terms* (e.g. “I,” “now,” “here,” and “that person there”), and (c) *definite descriptions* (“that x which has characteristic property F”). *Formal notation*: atomic singular terms are called *individual constants* and are formalized by small letters $a, a_1, a_2, \dots, b, \dots$

1.2. *First-order general concepts* are applied to singular terms: this application is called *predication* , and through it, an *atomic sentence* is generated.

1.2.1. *First-order predicates* denote features which can be true of individuals or not. One may distinguish between *property concepts* such as “x is big,” and *kind concepts* such as “x is a person.” *Formal notation*: predicates are formalized using capital letters $F, G, \dots$ (or $F_1, F_2, \dots$). Predicates are *one-place* , that is, they have one argument *place* , which is written as a *variable* after the predicate letter: “Fx” for “x is an F.” For example, if “Fx” stands for “x is slim,” and “a” for “Peter,” then “Fa” stands for the sentence “Peter is slim.”

1.2.2. A (first-order) *n* -place *relation sign* denotes a relation between *n* given individuals. *Formal notation* by uppercase letters $R, Q, \dots$, for example, if “Rxy” stands for the two-placed relation “x is a brother of y,” “a” for “Peter,” and “b” for “Paul,” then “Rab” stands for the sentence “Peter is a brother of Paul.”

1.2.3. *Function signs*: a one-place (or *n*-place) function sign denotes a function which uniquely assigns to one (or *n*) individuals a further individual. *Formal notation* usually by lowercase letters $f, g, \dots$. For example, if “f(x)” stands for “the mother of x” and “a” for “Peter,” then “f(a)” is a singular term designating Peter’s mother. In mathematics, functions play a fundamental role as operations on numbers; for example, addition + is a two-place function, which assigns to the respective numbers *x* and *y* their sum +(x,y), in infix notation written as “x+y.” Moreover, all *quantitative* (e.g. physical) *properties* of individuals are functions: for example, *m*(a) denotes the *mass* of the body *a*, and “*m*(a) = 5.6 kg” stands for “the mass of *a* is 5.6 kilograms.”

1.3. *Second (or higher)-order predicates and functions* are applied to sentences or predicates and thereby generate sentences. An example of a higher-order predicate is the concept of *truth* which applies to sentences.

2. *Logical concepts*: there are different kinds of logical concepts, and their principles are explicated in corresponding kinds of logics.

2.1. *Truth-functional sentence operators – Propositional logic PropL* (where *A, B, ...* stand for arbitrary propositions):

- the negation $\neg$ (read: $\neg A$ – *not A*),
- the conjunction $\wedge$ (read: $A \wedge B$ – *A and B*),
- the disjunction (the *inclusive* or) $\vee$ (read: $A \vee B$ – *A or B or both*),
- the so-called ‘material’ implication $\rightarrow$ (read: $A \rightarrow B$ – *if A, then B*),
- the ‘material’ equivalence $\leftrightarrow$ (read: $A \leftrightarrow B$ – *A if and only if (iff) B*),
- *convention*: “ $\leftrightarrow_{\text{def}}$ ” stands for “equivalent by definition”; similarly for “ $=_{\text{def}}$.”

Further sentence operators are thereby definable. For example, the exclusive *either-or* $\dot{\vee}$ is defined as $A \dot{\vee} B \leftrightarrow_{\text{def}} (A \vee B) \wedge \neg(A \wedge B)$ (*A or B, but not both*). The logical laws of these sentence operators are explicated by (non-modal) propositional logic.

A sentence operator, when applied to given sentences, generates a new, complex sentence. A sentence operator is called *truth functional* (or *extensional*) iff the truth *value* of the complex sentence is uniquely determined by the truth values of its component sentences (arguments). The truth-functionality of the sentence operators of PropL is expressed by their “truth tables”: every sentence *A* is either true or false; $\neg A$ is true iff *A* is false (else false); $A \wedge B$ is true iff both *A* and *B* are true (else false), etc. *Note*: the material implication $A \rightarrow B$ is defined as true iff *A* is false or *B* is true, that is, $A \rightarrow B \leftrightarrow \neg A \vee B$ is valid. The material implication is *weaker* than the natural language if-then, because its truth does not presuppose any connection between the contents of *A* and *B*.

2.2 *First-order quantifiers and variables – predicate logic PredL* : first-order *quantifiers* quantify over individuals with help of first-order *individual variables* $x, y, \dots (x_1, x_2, \dots)$. The two most important kinds of quantifiers are:

- the *universal quantifier* $\forall x$: read “ $\forall xFx$ ” as “For all x : Fx ,” and
- the *existential quantifier* $\exists x$: read “ $\exists xFx$ ” as “For at least one x : Fx .”

The class of individuals to which a quantifier refers is called its (semantic) domain D . So $\forall xFx$ says that all individuals in D have property F , and $\exists xFx$ that some individual in D has property F . When nothing further is stated, D is identified with the *universal* domain, that is, the class of all individuals in the given scientific area.

We say that the individual variable “ x ” is *bound* by the quantifier “ $\forall x$ ” or “ $\exists x$.” The subformula “ Fx ” in “ $\forall xFx$ ” is called the scope of the quantifier “ $\forall x$.” Formulas in which all individual variables are *bound* are called *closed formulas* or *sentences*. The expression “ Fx ” (for “ x is an F ”) is called an *open* formula, because the individual variable “ x ” is *free*, that is, not bound by any quantifier. Open formulas such as “ $Fx \wedge Gx$ ” express complex concepts.

A distinguished logical relation is *identity* “ $x = y$,” standing for “ x is identical with y ,” that is, “ x ” and “ y ” denote the same individual. The logic of sentences which are made up of singular terms, general first-order concepts, extensional sentence operators, and first-order quantifiers, is called first-order predicate logic PredL (introduction see [Klenk 1989](#) , [Copi et al. 2010](#)). First-order PredL contains PropL as a fragment and is one of the most important formal frameworks of philosophy of science, among other reasons because type-free set theory, upon which all of classical mathematics builds, is expressible in this framework.

2.3. *Intensional sentence operators* are not truth functional. An example is the necessity operator $\Box$ (“ $\Box A$ ” for “ A is necessary”): the truth value of “ $\Box A$ ” is not determined by the truth value of A . For instance, if A and B stand for the sentences “I have brown hair” and “ $2 + 2 = 4$,” respectively, then both A and B are true, but only $\Box B$ is true, while $\Box A$ is false, as the fact that I have brown hair is not necessary. Other examples are *deontic* sentence operators (“ A is obligatory”). The logic of intensional sentence operators is *generalized modal logic* (see [Hughes/Cresswell 1996](#)).

2.4. *Second (or higher) order quantifiers* quantify over states of affairs or properties and are dealt within second (or higher) order PredL. The logic of higher-order quantifiers is called logical *type theory*, which goes back to Russell ([1903](#)).

2.5. *Set-theoretical symbols*: sets are aggregations of individuals into a complex individual called a “set.” – *Formal notation*:

- “ $x \in y$ ” stands for “ x is an element of the set y .”
- $\{x: Fx\}$ stands for the set of all individuals x which possess the property F .
- $\{a_1, \dots, a_n\}$ stands for the set consisting of the individuals $a_1, \dots, a_n$.
- $(a_1, \dots, a_n)$ by contrast, stands for the ordered sequence consisting of $a_1, \dots, a_n$.

Note: sets are *invariant* with respect to (w.r.t.) permutations or repetitions of their elements, hence $\{a, b\} = \{a, a, b\} = \{b, a\} = \{b, a, a, b\}$ (etc.). By contrast, sequences or n -tuples consist in an ordering of individuals that allows for repetition: that is, $(a, b, b, c) \neq (a, b, c) \neq (b, a, c)$.

Set theory is a part of *informal* logic or metalogic, as “ $\in$ ” too is a logical concept in a broader sense. The axioms of type-free set theory (after *Zermelo* and *Fraenkel*) are expressed in the language of first-order PredL, but they go *beyond* purely logical principles (for an introduction, see van [Dalen et al. 1978](#) , [Hrbacek/Jech 1999](#)).

- “ $x \subset y$ (resp. $x \subseteq y$)” stands for “ x is a proper (resp. a proper or improper) subset of y .”
- “ $f: A \rightarrow B$ ” stands for “ f is a function which assigns to every individual x in A exactly one individual y in B ”.

Further set-theoretical notions are explained in the text.

2.6. *Mathematical concepts* , as with logical concepts, characterize *abstract-conceptual* objects or functions. They may be characterized either by internal axioms of mathematical theories or by set-theoretical definitions. Among the most important mathematical concepts are concepts for different kinds of numbers (e.g. $\mathbb{N}$ for the set of natural and $\mathbb{R}$ for the set of real numbers), and function concepts expressing operations on numbers (addition, multiplication, etc.).

3.1.2 Syntax and Semantics, Statements, and Models

Since Morris ([1946](#)), the dimensions of language are subdivided into *syntax* , *semantics* , and *pragmatics* . Syntax deals with the grammatical form of linguistic expressions, semantics with the meanings of expressions, and pragmatics finally, with those properties of expressions which are determined by their *context* . In this section, we outline some foundations of syntax and semantics.

From atomic concepts, one may construct terms and formulas of arbitrarily high complexity. The translation of a natural language sentence into a logical sentence is called *formalization*. Here are some examples:

(3.1-1) *Examples of formalizations:*

Translations: a – this animal, Rx – x is a raven, Bx – x is black.

- i. This animal is a raven and it is black. Formalization: $Ra \wedge Ba$.
- ii. All ravens are black. *Partial formalization:* For all x: if x is a raven, then x is black. *Formalization:* $\forall x(Rx \rightarrow Bx)$.
- iii. Some ravens are not black. *Partial formalization:* There exists an x which is a raven and is not black. *Formalization:* $\exists x(Rx \wedge \neg Bx)$.

Sentences with the same natural language syntax can have a very different logical syntax. In particular, the natural language copula “is” serves three different logical functions:

(3.1-2)

<i>Sentence</i>	<i>Formalization</i>	
	Ta	“is” as predication
i. Peter (a) is tall (T).	$a = f(b)$	“is” as identity
ii. Peter is the father (f) of Paul (b).	$\forall x(Rx \rightarrow Bx)$	“is” as inclusion of properties
iii. A raven (R) is a bird (B).		

In ([3.1-2](#))(iii), there is a *concealed* universal quantifier, which is revealed by formalization. The formalization of ([3.1-3](#)) reveals a *concealed* existential quantifier:

([3.1-3](#)) Peter (a) has a son. *Partial formalization:* There exists an x, such that x is the son (S) of Peter. *Formalization:* $\exists x Sxa$.

Logical formalization is an important aid to the advanced philosophy of science, to uncover the logical structure of sentences, and building on this, to discover logical connections. In this book, formalization is used rarely and as little as possible; for most purposes, partial formalizations such as in ([3.1-1](#)) above are sufficient.

In *semantics* , one differentiates between two kinds of semantic denotation: every expression has on the one hand a *meaning* , and on the other an object *reference* . (Terminologies differ, however; cf. [Kutschera 1976](#) , ch. 2; [Morris 2006](#) .) The meaning or intension of a concept is that which must be known to a speaker in order to *understand* the concept properly. The meaning of an expression *together* with the *facts* of the world determines the expression’s reference, that is, the real entity denoted by the expression (if there is no such entity, the reference is empty). There are controversial views on the question of wherein *exactly*

meaning exists (in the imagination, in thought, in the definition, or in usage?). The reference of a singular term is uncontroversially an individual. The reference of predicates is also controversial: some authors (e.g. [Carnap 1956b](#), ch. I) identify the reference of a predicate with its *extension*, that is, the *set* of individuals which fall under the predicate; other authors (e.g. [Fodor 1990](#), 93) identify the reference of a predicate with the *real property* or natural kind denoted by it.

The meaning of sentences is usually identified with the *proposition* expressed by the sentence. According to the traditional view, logically or analytically equivalent sentences express the *same* proposition. The reference of sentences is sometimes identified with truth values (e.g. in [Frege 1892](#)), and sometimes with events or real states of affairs. The following definition of “state of affairs” and “fact” is uncontroversial: every descriptive sentence expresses a *state of affairs*; and if the sentence is true, the state of affairs really exists and is then a *fact*.

In *logical semantics*, which goes back to *Alfred Tarski* (1936a), one assigns to linguistic expressions formal object references, in the form of set-theoretical *interpretations*. Informal set theory functions in this case as a *metalanguage*, in which the semantic relationship between the *object language* and a hypothetical object structure is modeled. The concrete definition is as follows: an extensional *interpretation* of a logical language $\mathcal{L}$ is a pair (D, I) , consisting of a domain of individuals D and an interpretation function I , which assigns to every individual constant a and variable x from $\mathcal{L}$ an individual $I(a)$ in D , to every n -placed relation sign R in $\mathcal{L}$ a subset $I(R)$ from D^n (where D^n is the set of all n -tuples of D -individuals), and to every n -placed function sign f in $\mathcal{L}$ a function $I(f): D^n \rightarrow D$. An extensional interpretation (D, I) makes an atomic formula Fa true, in short $(D, I) \models Fa$, iff $I(a) \in I(F)$. This basic truth definition is recursively extended to arbitrary formulas as usual (cf. [Machover 1996](#), sec. 8.4; [Bergmann et al. 1998](#), sec. 8.7).

The extensional interpretation of an object language can also be seen as a *relationship of translation* between a logical and a set-theoretical language – for example as follows:

(Prop. 3.1-1) *Translation (I) between logical and set-theoretical language*, where $I(a) = a$, $I(F) = F$, etc.:

Logical sentence	Set-theoretical sentence (translation)
Fa	$a \in F$
$\forall x(Fx \rightarrow Gx)$	$\forall x(x \in F \rightarrow x \in G)$, in short: $F \subseteq G$
$\psi(F)$	$F \in \psi$ (etc.)

In set-theoretical translations, then, individual constants and predicates or relation symbols are replaced by set terms which denote their extensions, and the element relationship $\in$ is inserted in between. This procedure is iterable for higher-order predications. Set-theoretical formalizations may be less intuitive than logical formalizations, but they have the advantage to be *type free*, that is, they allow for the translation of all sentences of (higher-order) type theory into the language of first-order PredL. For this reason, structuralist philosophers of science have argued that set-theoretical formalizations are preferable to logical formalizations (see [sec. 1.3.6](#)).

An extensional interpretation (D, I) of a language $\mathcal{L}$ is also called a *model* for $\mathcal{L}$. It plays the role of a “possible world” as expressible by means of $\mathcal{L}$. Often the models (D, I) of a sentence are written in the form $(D; a_1, \dots, a_n; R_1, \dots, R_m; f_1, \dots, f_k)$, in which instead of I , the designata of all individual constants, relational and function symbols assigned by I are listed. In this form, models are used by proponents of the (German) *structuralist* view and the (Anglo-Saxon) *semantic* view of theories. Here, a theory is represented by the set of its models, $M(T)$ – a method that goes back to Suppes ([1999](#), 249). By contrast, in the traditional “statement” view, a theory is represented by the set of its axioms, $Ax(T)$, together with semantic characterizations of its concepts.

The (traditional) “statement view” and the (structuralist) “non-statement view” have been set in opposition by Stegmüller ([1976](#)). This “opposition” is largely unjustified, for two reasons. *First*, the statement view is not (as is sometimes claimed) only a purely syntactical, but also a semantical view,

inasmuch as the statements of $\mathcal{L}$ are taken as semantically interpreted propositions (cf. [French 2008](#) , 272). Since his turn in 1947, Carnap understood a linguistic framework always as a language $\mathcal{L}$ *together* with a semantic interpretation of $\mathcal{L}$ ([Carnap 1956b](#)). *Second* , model-theoretic theory reconstructions are not free from using statements (which is wrongly suggested by the name “non-statement view”). Rather, the class of models $M(T)$ of a theory T is defined by nothing but the set of its linguistic axioms $Ax(T)$, as follows: $M(T) =_{\text{def}} \{m \in M(\mathcal{L}) : m \models Ax(T)\}$, with $M(\mathcal{L})$ as the set of models for the language $\mathcal{L}$ in which T is formulated (cf. [Balzer et al. 1987](#) , ch. 1). Therefore, there exist straightforward *translation relations* between linguistic and structuralist theory-representations. Here is a simple example:

(Prop. 3.1-2) *Linguistic and model-theoretic theory representation* (simplest case):

Linguistically by T's axioms:	Model-theoretically by T's models:
Semantic ‘axioms’:	Models: $M(T)$ = the class of models
$A_1 : I(F) \subseteq D. A_2 : I(G) \subseteq D.$	(D,F,G) such that $(D,F,G) \models \{A_1, A_2, A_3\}$.
Factual axiom $A_3 : \forall x(Fx \rightarrow Gx).$	

Often translation relations are more complicated, since structuralist theories contain several ingredients; for details see Andreas ([2010](#)) and Schurz ([2014b](#)).

3.1.3 The Content Type of Concepts

This classification is particularly significant in the philosophy of science. It applies only to *nonlogical concepts* . [Fig. 3.1-2](#) shows an *overview* .

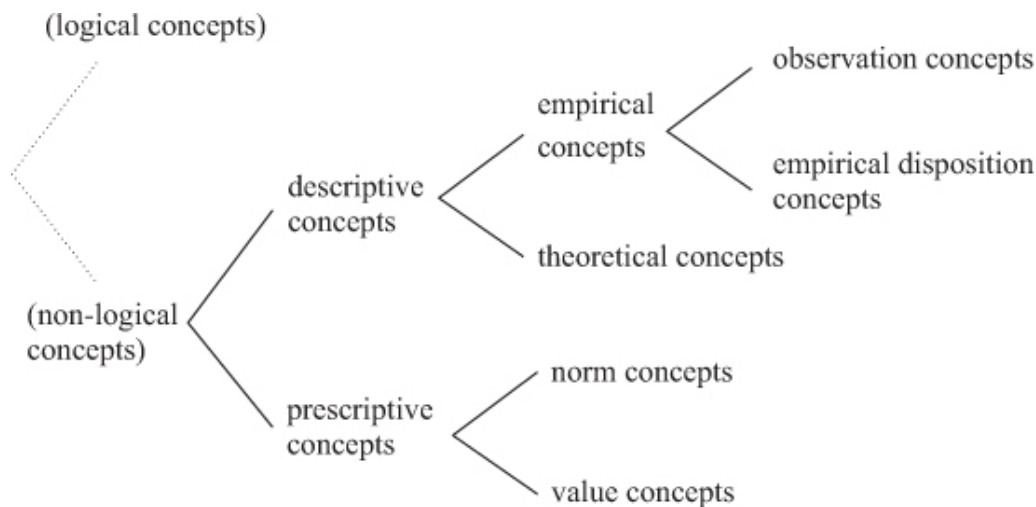


Figure 3.1-2 Classification of concepts according to their content type. ↵

First, we have to distinguish between descriptive and prescriptive concepts. *Prescriptive* concepts consist of norm and value concepts ([Hare 1952](#)). There are only a few prescriptive concepts: the *normative* concepts of something being obligatory, forbidden, or permitted, and the qualitative, comparative, and quantitative *value* concepts (X is ethically or aesthetically valuable, or more valuable than Y , or valuable to a certain degree). Philosophy of science deals primarily with *descriptive* concepts and sentences which describe the facts of the world. There is an enormous number of (singular and general) descriptive concepts of various types.

The *most important* distinction within the descriptive concepts is the one between empirical and theoretical concepts, and within the empirical concepts the one between observation concepts and empirical dispositional concepts.

(1.) A general *observation concept in the narrow sense* (i.n.s.) expresses a property that is perceptible to the senses, or a complex of such properties. As we saw in [sec. 2.9.2](#) , a property perceptible to the senses is one which can be *ostensively learned* by humans, independently of their background knowledge. *Examples*: “x is red,” “x is a raven,” “x is longer than y,” “the needle of the measuring instrument x is pointing to 5.” The classical and the early logical empiricists demanded that all descriptive concepts be observational concepts. In fact, only a *few* concepts are genuine observation concepts in this narrow sense.

Note the following subtlety: strictly speaking, a property as such is not perceptible, but only a certain *instantiation* of a property (a so-called *trope*) in a perceptible individual. “Being red,” for example, expresses an observational property as applied to visible objects; but in application to non-observable individuals such as microbes or molecules, the property of being red is not observable, as these individuals are too small to be visible. For this reason, Quine was right when he stressed that the primary level of “observability” is not that of concepts, but of “occasion sentences” like “this here is red” (1960, 35ff; 1992, 3). Putnam ([1962](#) , 242f) saw in this fact a fundamental obstacle to a clear definition of observation concepts. Our criterion of ostensive learnability, however, solves this problem as follows: a primitive property is ostensively learnable, and thus observable i.n.s., iff it possesses at least *some* perceptible instantiations in perceptible individuals which may serve as training instances.

(2.) An *empirical dispositional concept* expresses an empirical disposition. By this, we understand the *lawlike regularity* of an individual to produce, under certain empirical circumstances or trigger conditions, a certain observable *result* that is also called the disposition’s “manifestation” ([Schrenk 2017](#) , 45). Example: the empirical dispositional concept “x is soluble in water” is by definition equivalent to the following lawlike hypothesis about “x” which generalizes over time points: “whenever x is or were placed in water, it will or would dissolve.” Lawlikeness and counterfactual conditionals (“would”) go beyond the observable and presuppose inductive projectibility (see [sec. 3.6](#) , 6.5.3). Nevertheless, dispositional concepts are *empirically operationalized* concepts, whose meaning is not dependent on a theory. Further examples: “x is hard,” “x weighs 4 kg on a pair of scales,” “person x is a German speaker.” Empirical dispositional concepts correspond to what the *operationalists* (e.g. [Bridgeman 1936](#)) demanded of *all* scientific concepts. In fact, though, theoretical concepts are not reducible to empirical dispositional concepts (see [sec. 3.10](#)).

(3.) An *empirical concept i.n.s.* is an observation concept or an empirical dispositional concept.

(4.) A *theoretical concept in the wide sense* (i.w.s.) is any descriptive concept which is not an empirical concept i.n.s., but still has a meaning. A theoretical concept i.w.s. does not have to be part of a scientific theory; it can also be based on speculation (e.g. “the prime cause of all being”).

(5.) A *theoretical concept i.n.s.* is a theoretical concept i.w.s., whose meaning is *introduced* by a scientific theory. If we want to make explicit the reference to a specific theory T, we speak of a *T-theoretical concept* ([Sneed 1971](#) ; refinements in [sec. 5.3.2](#)). *Examples*: physical force, kinds of molecules, biological genotype, psychological character, intention, and social order.

Convention: unless otherwise stated, we understand empirical concepts and theoretical concepts respectively in the narrow sense (i.n.s.).

(6.) As discussed in [sec. 2.9.1](#) (pro-5), a number of philosophers of science use the notion of an empirical concept in the wide sense (i.w.s.), meaning any concept that expresses a *measurable property* or ‘datum.’ Empirical concepts i.w.s. are theory dependent – their meaning depends on theories which describe the functioning of the measuring instruments and are presupposed to be unproblematic. The border between empirical concepts i.w.s. and theoretical concepts i.n.s. can only be drawn *pragmatically* and in a *context-dependent* way. By contrast, there exists an *objective* but gradual border between empirical concepts i.n.s. and theoretical concepts i.w.s.

3.1.4 The Gradation (Scale) Type of Concepts

Most properties in our world do not only occur or not occur; they come in certain *degrees* . The gradation or scale type of a concept reflects how finely the concept is graduated. Accordingly, scientific concepts may be classified as in the *overview* in [fig. 3.1-3](#) .

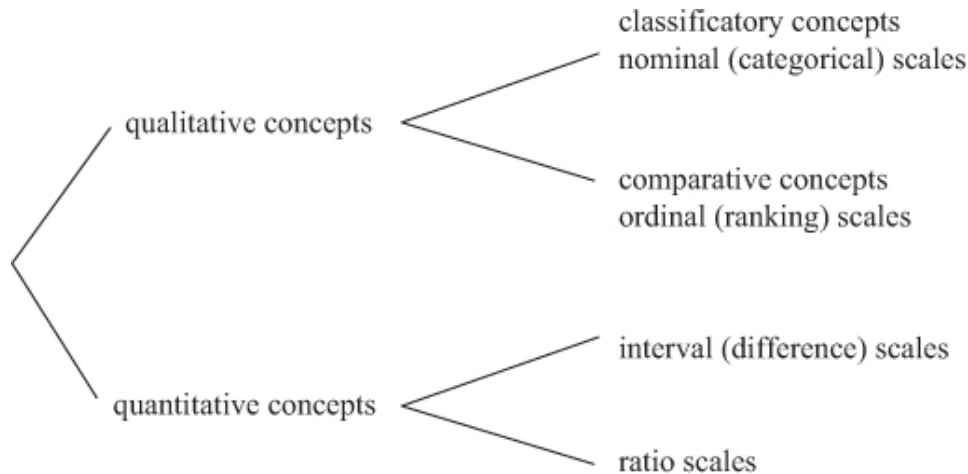


Figure 3.1-3 Classification of concepts according to their gradation type. [↗](#)

Qualitative concepts are not graduated, or more precisely, they are graduated only in the *yes-no* sense; comparative concepts are graduated in the sense of an *ordering*, and quantitative concepts are the most strongly graduated, that is, in a *continuous* way. In science, the graduation of a property is measured in terms of numbers, and we speak of the scale type of concepts (e.g. [Krantz et al. 2006](#), ch. 1; [Bortz 1985](#), 28–33). The *finer* graduated a concept is, the *sharper* the information we express with it, and the *stricter* the tests to which the hypotheses using it can be subjected. The types of gradation will now be discussed.

3.1.4.1 Qualitative-Classificatory Concepts and Nominal Scales

With the help of qualitative or classificatory concepts, we divide a domain D of objects into certain subgroups $D_1, \dots, D_n$, which are defined by their having certain properties or belonging to certain categories $C_1, \dots, C_n$. Here, it is demanded that: (a) the classification is *disjoint*, that is, the categories C_i must be mutually exclusive (the subgroups must not overlap), and (b) the classification is *exhaustive*, that is, the union of the subgroups must contain all of D .

The set $F =_{\text{def}} \{C_1, \dots, C_n\}$ of the categories of a classification is called a (classificatory) *concept family* ([Carnap 1971](#), 43ff). In the simplest case, an analytically independent concept Fx forms a two-valued or *binary* concept family $\{Fx, \neg Fx\}$: for example, x is either human or not. The conditions of disjunctivity and exhaustivity hold in this case purely logically. There are also concept families with more than two categories; for example, the family of color concepts $\{\text{red, yellow, green, and blue, } \dots\}$, or the family of social classes $\{\text{lower class, middle class, and upper class}\}$. The conditions of disjunctivity and exhaustivity are in these cases based on analytical conventions.

Every concept family F corresponds to a family concept or *attribute* X_F . For example, the concept family $\{\text{lower, middle, and upper class}\}$ corresponds to the attribute “social class of x .” Formally, X_F is a higher-order *function*, which assigns to every object in D its characteristic category, that is, $X_F : D \rightarrow \{C_1, \dots, C_n\}$. If we replace the categories C_i by *natural numbers* i , for example, “lower, middle, upper class” by “1,2,3,” then we obtain a nominal (or categorical) *scale*. For nominal scales, however, the use of numbering is inessential and arbitrary. In mathematics, a function X which when applied to various objects can take on various *numerical values* is also called a *variable*. Simple binary properties (F) correspond to a *binary* variable X_F with values F and $\neg F$. Observe that mathematical variables are functions; they should not be confused with variables in the logical sense.

3.1.4.2 Comparative concepts and ordinal scales

With the help of comparative concepts, the objects of a domain D are *ordered* according to the grade of a given property, that is, they are divided into *ranking groups* $D_1, \dots, D_n$ (see [fig. 3.1-4](#)): D_1 contains all D -objects with the highest grade of the given property, D_2 all D -objects with the next highest grade, etc. The

objects (“o”) within a rank group have the same grade. If the rank groups D_i are assigned natural numbers i , we have an *ordinal* or *ranking scale*. The corresponding assignment function $r: D \rightarrow \{1, \dots, n\}$ is an ordinally scaled variable, also called a ranking function.

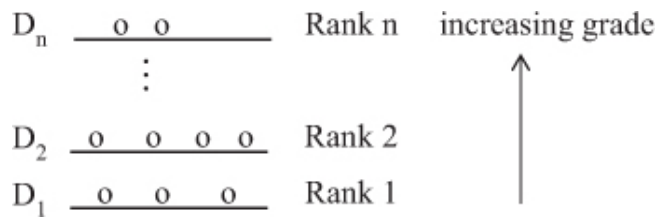


Figure 3.1-4 Ordinal scale of D. [↗](#)

The ranks of an ordinal scale give us only comparative information; quantitative differences between the ordinal ranks are unknown or undefined. In the *Mohs scale*, for example, the hardness of minerals is determined by a reciprocal *scratch comparison*: if mineral A scratches mineral B, but not vice versa, then A has a higher degree of hardness. Mohs used this method to devise a ten-point hardness scale: 1 (talcum), 2 (gypsum), 3 (calcium carbonate), up to 9 (corundum), and 10 (diamond). We must not conclude from this that the difference in hardness between talcum and gypsum is for example one-ninth of the difference between talcum and diamond. Similarly, the intensity of *psychological* attitudes can typically be measured only by an ordinal scale: for example, I can find one dish *tastier* than another without being able to say how much *tastier* it is.

An important method used in the empirical social sciences for *attitude surveys* consists in asking test persons (tps) to place their opinions on a given topic into a *preset* ordinal scale. A questionnaire on the attitudes of tps toward old age could for example contain the following question:

Please tick where appropriate: I believe that securing an old age pension is:

- ☐ very important ☐ quite important ☐ don't know
☐ not so important ☐ not important at all

The attitudes of the persons surveyed are thereby measured by an ordinal scale with five ranking levels, from not important at all (1) up to very important (5).

The basic logical concept of ordinal scales is that of a *smaller-or-equal* relation $\leq_P$ in regard to the given property P : $x \leq_P y$ states that the P -grade of x is less than or equal to that of y . With help of “ $\leq_P$,” one defines the relation of *P-equality* as $x \equiv_P y \leftrightarrow_{\text{def}} x \leq_P y \wedge y \leq_P x$ (x 's P -grade equals y 's), and the *strictly smaller* relation as $x <_P y \leftrightarrow_{\text{def}} x \leq_P y \wedge \neg y \leq_P x$ (x 's P -grade is smaller than y 's). The concepts $\leq_P$, $\equiv_P$, and $<_P$ are operationalized by *empirical* operations of comparison. In order to generate a *unique* ranking scale, these comparison operations must fulfill at least the axioms of a total *quasi-order*: *reflexivity* ($\forall x: x \leq_P x$), *transitivity* ($\forall x, y, z: x \leq_P y \wedge y \leq_P z \rightarrow x \leq_P z$), and *connexivity* ($\forall x, y: x \leq_P y \vee y \leq_P x$), that is, everything is comparable to everything else. Axioms and definitions entail the *anti-symmetry* of $\leq_P$ ($\forall x, y: x \leq_P y \wedge \neg x \equiv_P y \rightarrow \neg y \leq_P x$) and the fact that $\equiv_P$ is an *equivalence relation*, that is, fulfills the axioms of reflexivity, symmetry, and transitivity. (Note: an *order* in the mathematical sense arises from a quasi-order if no two D -objects occupy the same rank.)

The axioms of transitivity and connexivity are prerequisites for ordinal scaling. Whether they are at least nearly fulfilled is an *empirical* question (only reflexivity and anti-symmetry are fulfilled analytically). It is well known that the empirical preferences of tps are not always transitive (cf. [Garnham/Oakhill 1994](#), 184ff). In addition, all empirical *pair comparisons* have a certain *inaccuracy*. The relation of *approximate equality*, though, is no longer transitive when the grade differences between objects are very small. As an example, assume that a is 0.6 mm taller than b , b is 0.6 mm taller than c , but that only height differences of

more than 1 mm are observable. Then we observe that a is approximately equal to b, and b to c, but a is taller than c. If we place objects of this sort in a *preset* ordinal scale, we make their rank relations artificially transitive: the result may then depend on the order in which the objects are considered by the tp.

3.1.4.3 . Quantitative concepts and scales

These express the grade of a property P by means of a *numerical value* , along with a *unit of measurement* . Logically speaking, a quantitative concept is a function $f(x)$, which assigns to every object x in domain D a numeric value $r \in \text{Num}$ along with a unit of measurement e_f ; for example, $\text{mass}(x) = 5 \text{ kg}$, or $\text{temperature}(x) = 20^\circ \text{ Celsius}$. If $f(x)$ is a *discrete* magnitude, then f assigns integers to the objects; if $f(x)$ is a *continuous* magnitude, then f assigns real numbers to the objects. Almost all physical magnitudes are continuous; exceptions occur in quantum mechanics. Typical discrete magnitudes are counting numbers, for example, population sizes. In metaphysics, the nonspecified magnitude (e.g. “mass”) is also referred to as a determinable, and its numerical instantiation (e.g. “mass of 3.791 kg”) as a determinate property ([Wilson 2023](#)).

Directly observable are at best *comparative* magnitude relations: if we determine the length of a given distance by using a measuring tape, we can see whether this distance is longer, shorter, or approximately equal to certain scale divisions on the measuring tape. Converting a comparative concept into a quantitative concept by introducing an (empirical) measuring operation is called its *metrization* . In order to make a comparative property quantitatively measurable, certain empirical conditions must be satisfied, which are explicated axiomatically in *metrization theory* (see [Krantz et al. 2006](#)). On the other hand, any empirical measurement operation is based on certain *arbitrary conventions* , to which the results of respective measurements are relativized. Interval and ratio scales differ from each other in the amount of these arbitrary conventions. In an *interval scale* , both the *zero point* of the scale and the *unit of measurement* are set by convention. In a ratio scale, on the other hand, only the unit of measurement is conventional, while the zero point is empirically determined.

The empirical temperature scale, for example, is an interval scale. It is based on the fact that certain liquids, for example, mercury in a glass column, expand uniformly as the temperature increases. But there is no zero point of warmth or coldness which might be discerned through mere observation. In the *Celsius scale* ($^\circ \text{C}$), the freezing point of water is used as a conventional zero point, and the scale unit of 1°C is conventionally defined as one 100th of the length of expansion of mercury in a glass column as it rises from the freezing point to the boiling point of water. In the *Fahrenheit-scale* ($^\circ \text{F}$), the freezing point of water is set at 32°F , and the scale unit of 1°F corresponds to $5/9^\circ \text{C}$. So physically, it makes no sense to say that a liquid at 20°C is *twice as warm* as one at 10°C . This relationship depends on the artificial choice of the zero point; in the $^\circ \text{F}$ scale, the same relationship comes out as 68°F to 50°F . If, however, we have three liquids; a, b, and c at 10°C , 20°C , and 30°C , respectively, then it certainly does make sense in physical terms to say that the *temperature difference* between b and a is exactly as great as the one between c and b, and half as great as the one between a and c, as this is valid *independently* of the choice of zero point: in the difference between two temperature values, the zero point is canceled out. Because only statements about ratios of temperature intervals are objective, while statements about ratios of temperatures lack objective content, we talk of interval scales. Measurements of *position* and *time* also use interval scales, because the zero points of spatial and temporal coordinate systems are set by pure convention. We cannot say for example “the year 2000 is twice as late as the year 1000,” as this is relative to the birth of Christ as the starting point of the calendar. If our calendar began in the Middle Stone Age, we would have instead the years 12000 and 11000.

In *ratio scales* , unlike interval scales, the absolute zero point is empirically determined. For example, mass, volume, length (as opposed to spatial position), and duration (as opposed to temporal position) are ratio-scaled magnitudes. Physically, it makes sense to say that Peter, at 100kg, weighs twice as much as his 50 kg little brother. Because such ratio statements have objective content, we speak of ratio scales. Similarly, quantitative magnitudes in economics, such as amounts of money or goods, are based on ratio scales.

These magnitudes are also called *extensive* magnitudes, because when joining or “concatenating” two objects (a, b) into a whole (a o b), the resulting magnitude grows monotonically, that is, it holds that $f(a \text{ o } b) > f(a), f(b)$. ¹ Magnitudes such as temperature, for which this is not the case, are by contrast called *intensive* magnitudes – the temperature of say, water, is not increased if we increase the amount of water.

The metrization of extensive magnitudes leads to ratio scales. What is still arbitrary in ratio scales is the (conventional) choice of a *unit*. It makes no sense to say that the numerical value of Peter's weight is 100 – this is only valid if the unit of mass is assumed to be the kilogram; if we chose the gram as unit instead, the numerical value of Peter's mass would be 100,000. One subtle problem is the justification of the assumption that the chosen unit remains *constant over time* (see on this [Carnap 1966](#), ch. 8–9).

Only in simple *counting scales* is the unit naturally determined as one individual or item; such counting scales are also called *absolute scales*. The difference between various types of scales is mathematically described in terms of *admissible scale transformations*. The coarser a scale, the more scale transformations are admissible; for details see Krantz et al. ([2006](#), 11).

In theories of metrization, one proves, for example, the following: an extensive comparative magnitude relation $\leq_p$, for example, “is longer than,” over a domain of objects D is metrizable in the form of a ratio scale iff $\leq_p$ is a monotone quasi-order over D which satisfies the so-called *Archimedean* condition (see [Krantz et al. 2006](#), 73). The latter condition demands that every object b, no matter how large, can be balanced out by a sufficient number of copies of an arbitrary object a, no matter how small a is. For social preference relations, the Archimedean condition is often not satisfied – for example, one would be reluctant to say that the value of a human life might be balanced out by a sufficient number of banknotes (cf. [Schurz 1997](#), sec. 11.5.2).

Strictly speaking, the results of attitude measurements (e.g. 5 = very important, 3 = don't know, and 1 = unimportant) only have the level of an ordinal scale, or they lie somewhere “between” ordinal- and interval-scale levels. In spite of this, they are often evaluated using statistical tests which presuppose interval scales. To assess whether such an idealization is harmless in a given case, it is necessary to reevaluate the results of interval scale-based tests by an ordinal scale-based (or “non-parametric”) test (see [Bortz 1985](#), 31f, 178f).

So far, we have been concerned with *fundamental empirical* metrization. One speaks of *derived* metrization when the metrization of a concept is traced back to the metrization of other concepts, for example, density to the quotient of mass divided by volume. One speaks, finally, of *theory-dependent* metrization, when the metrization is based on a certain theory, for example, from physics ([Balzer 1985b](#), 6). In theory-dependent metrizations, a change in the scale level of an initially empirically defined metric may occur for theoretical reasons. An example of this is the *absolute* temperature scale after Kelvin ($^{\circ}\text{K}$), which postulates the absolute zero point of temperature at $-273.15^{\circ}\text{C} = 0^{\circ}\text{K}$. This assumption resulted from the *kinetic theory of gases*, according to which heat and temperature are to be attributed to the movement of molecules. In ideal gases, the volume (V) increases in a *linear* way with rising temperature ($V = k \cdot T$). If one extrapolates this relationship in a T×V-coordinate system to the left into the region of negative $^{\circ}\text{C}$ -temperatures, one reaches a temperature of -273.15°C , at which the volume of an ideal gas would have to be *zero*: here, the gas molecules would be crammed into an infinitely small space, and thus would be unable to move. Kelvin therefore identified this point as the absolute zero of temperature. His hypothesis that no cooling process could go below this absolute zero point has since been confirmed many times.

3.2 Classification of Sentences According to Their Content Type

Sentences may also be classified into several different dimensions. Most significant in philosophy of science is the classification of sentences according to their content type, which is treated of here, and according to their degree of generality, which is dealt with in sec 3.5. [Fig. 3.2-1](#) gives an *overview*.

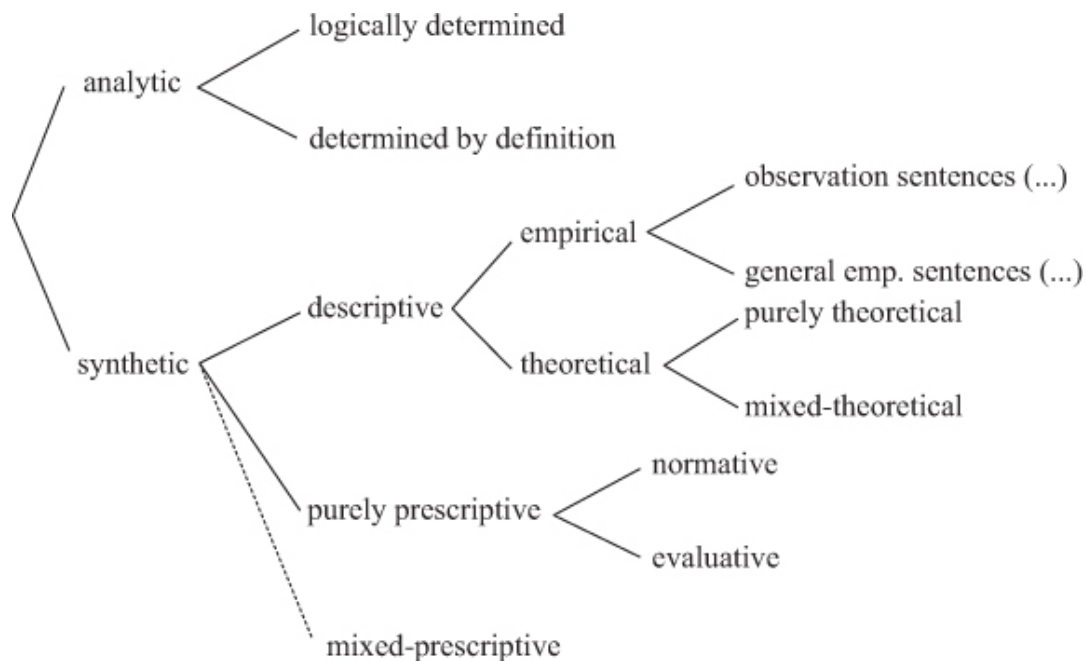


Figure 3.2-1 Classification of sentences according to their content type. [↗](#)

Analytically true sentences are those whose truth is determined by the laws of logic or by extralogical conventions of linguistic meaning, independently of the factual constitution of the world. Analytically false sentences are, by analogy, those whose falsity is determined in this way, and *analytic sentences* are those whose truth value (true or false) is determined in this way. Analytic sentences may be broken down further into logically determined sentences and sentences determined by definition or extralogical meaning convention ([sec. 3.3](#) – [3.4](#)).

By contrast, *synthetic* sentences are those sentences whose truth value is not analytically determined, but is dependent on the factual constitution of the world, or in the case of prescriptive-synthetic sentences, on our moral interpretation thereof. As with our content-based classification of concepts, we first divide synthetic sentences into descriptive and prescriptive ones; more precisely, into purely descriptive sentences, purely prescriptive sentences, and mixed sentences with descriptive and prescriptive components. Descriptive sentences tell us something about the factual composition of the world; prescriptive sentences on the other hand express propositions about norms or values. In the following subsections, we refine these distinctions.

3.2.1 Descriptive versus Prescriptive Sentences

Not every sentence containing a prescriptive concept is thereby prescriptive. Indeed, whenever a prescriptive concept lies within the *scope* of a *subjective attitude operator* – that is, a sentence operator expressing a subjective attitude toward a state of affairs – the prescriptive concept loses its prescriptive force. So for example, the sentences

Peter believes that stealing is wrong, – or
 Peter believes that you shouldn't steal,

are empirically descriptive, as they state something about Peter's factual attitude to values or norms. The same goes for all subjective attitude operators, like for example: person X believes, thinks, feels, says, wishes,..., that A. The significance of this fact for the value-neutrality thesis was worked out in [sec. 2.5.1](#) . We define:

(Def. 3.2-1) (a) A sentence is *descriptive* iff every prescriptive symbol occurring in it lies within the scope of a subjective attitude operator.

(b) A sentence is *purely prescriptive* iff all of its descriptive components lie within the scope of a prescriptive operator.

(c) A sentence which is neither descriptive nor purely prescriptive is called *mixed-prescriptive*. These sentences contain both descriptive components outside the scope of any prescriptive operator and purely prescriptive components outside the scope of any subjective attitude operator.

(3.2-1) *Examples of descriptive sentences* (logical structure at the right; prescriptive operators underlined):

- i. Peter steals. – Steals(Peter).
- ii. Peter believes that stealing is forbidden – Believes(Peter, Forbidden (stealing)).

Examples of purely prescriptive sentences:

- iii Stealing is forbidden. – Forbidden (stealing).
- iv If an action is under all circumstances morally bad, then it is forbidden. – $\forall a,c$ (Bad (a,c) $\rightarrow$ Forbidden (a)).

Examples of mixed-prescriptive sentences:

- v For persons suffering from hunger, stealing is allowed. – $\forall x$ (Suffering-hunger(x) $\rightarrow$ Permitted (Steals(x))).

Two remarks: (a) implicational connections between purely prescriptive sentences as in iv are themselves purely prescriptive. (b) Def. 3.2-1 is a *syntactic* definition; it can be refined by the method of *essential* sentence types in [def. 3.5-1](#).

3.2.2 Observation and Basic Sentences, Empirical and Theoretical Sentences

Building on the content-based classification of concepts in [sec. 3.1.3](#), we classify descriptive sentences according to their content type. Here, we have to address two subtleties which are rarely discussed in the literature. They have to do with the fact mentioned in [sec. 3.1.3](#), that instantiations of observable properties are only observable if the individuals possessing them are also observable.

First subtlety: an *empirical* sentence is usually defined as a sentence which (apart from logical concepts) contains only empirical concepts (cf. [Tuomela 1973](#), 23f; [Kutschera 1972](#), 260, 292). But this is insufficient. Consider a universal sentence “All Fs are Gs,” which contains only the observation concepts F and G. This universal sentence can only be called empirical if it refers solely to perceptible individuals. Whether this is the case depends on the concept F: only if it follows from the meaning of F that F can be applied *only* to perceptible individuals is the universal sentence an empirical sentence. For example, “all ravens are black” is a universal empirical sentence, because it follows from the meaning of “raven” that only perceptible individuals can be ravens. However, “all loose objects on Earth fall downwards” is not a universal empirical sentence if by “objects” we are also talking about microscopically small particles. The same applies to existential sentences. Lewis ([1970](#), 79) and van Fraassen ([1980](#), 13f) concluded from this problem that one cannot ascertain the empirical versus theoretical nature of a sentence from the meaning of its concepts. But as we will show now, this is indeed possible. For this purpose, we introduce the following *definition*: a quantifier in a given sentence has an *empirical range* iff this sentence is analytically equivalent with its quantificational restriction to observable individuals only. For a universal sentence $\forall xA$, this restriction is defined as the sentence $\forall x(Ox \rightarrow A)$, and for an existential sentence $\exists xA$ as $\exists x(Ox \wedge A)$, where “ Ox ” stands for “ x is observable.” Otherwise, the quantifier has a *theoretical range*.

Second subtlety: observation sentences talk only of a finite or manageable number of individuals, because what humans can observe is of finite and limited complexity. It would be inadequate to limit observation sentences to *singular sentences* – that is, sentences *without* quantifiers – which contain only observation concepts. *Localized existential sentences* of the form “at point k in space-time, there is such-and-such an observable thing” are also observation sentences; for example, “in this basket (now) there is a red apple” ([Popper 1935](#) /2002, sec. V.28, speaks of “singular existential sentences”). This is the case because the

existence quantifier in such a sentence has empirical range, and because the limited segment of space-time is small enough to afford us a visual overview. For similar reasons, even *localized universal sentences* have to be counted as observation sentences, which have the following form: all (observable) individuals in this limited area of space-time have this-or-that observable property”; for example, “All apples in this basket are red.” Once again, the universal quantifier must have an empirical range; in addition, it is important that the individuals satisfying the antecedent predicate have a perceptible spatial extension that does *not overlap* with other individuals. Given these conditions, it follows from the meaning of “apple” and “basket” that there is only room for a finite and manageable number of apples in the described basket, so that we can determine the truth value of the sentence by observation.

This brings us to the following definition:

(Def. 3.2-2) If S is a synthetic-descriptive sentence, the following holds:

- a. S is an *observation sentence* iff apart from logical concepts, S contains only observation concepts, and S is either a singular sentence, or a *localized* existential or universal sentence, whose quantifiers possess empirical range.
- b. S is an *empirical sentence* (i.n.s.) iff apart from logical-mathematical concepts, S contains only empirical concepts i.n.s., and its quantifiers possess empirical range.
- c. S is a *theoretical sentence* i.w.s. iff S contains theoretical concepts i.w.s. or its quantifiers possess theoretical range.
- d. S is a *T-theoretical* sentence (or theoretical i.n.s.) iff S is a theoretical sentence i.w.s. and the theoretical concepts of S are introduced by the scientific theory T (refinements in [sec. 5.3.2](#)).

In analogy to [sec. 3.1.3](#) , we speak of an empirical sentence i.w.s. if apart from logical-mathematical concepts, it contains only empirical concepts i.w.s. (and quantifiers with an empirical range i.w.s.). Empirical sentences i.w.s. are dependent on unproblematic pre-theories and reflect what scientists pragmatically call ‘data.’ *Convention:* unless otherwise stated, we understand observation, empirical, and theoretical sentences i.n.s.

The class of empirical sentences is much more *comprehensive* than the class of observation sentences. Apart from observation sentences, it also comprises: (i) *universal* empirical sentences like for instance “all ravens are black,” (ii) singular empirical *disposition* sentences like “this is soluble in water,” and (iii) general empirical disposition sentences like “all sugar is soluble in water.” An important subgroup of empirical disposition sentences are empirical *measurement* sentences like “this tree is (rounded off) 10.52 meters high,” or “fully grown fir trees are more than 10 meters high.”

Theoretical hypotheses need not always be universal sentences. Singular or existence sentences can also be theoretical. For example, the singular sentence “this mineral is radioactive” expresses a theoretical hypothesis which can only be tested by means of theory-led measurements using complicated instruments. Finally, *theories* are *systems* of theoretical sentences (ch. 5). We stress here that *theories* or theoretical sentences may *also* contain empirical concepts, and even *must* do so in order to have empirical content, because some authors (e.g. [Putnam 1962](#) ; [Chalmers 1986](#) , 36; [Quine 1992](#) , 6–8) give the misleading impression that a concept, once it occurs within a theory, becomes a theoretical concept. According to our definitions, this is clearly not the case (recall [sec. 2.9.1](#) , contra-7).

Narrower still than the category of observation sentences is the category of *basic sentences* : basic sentences in the sense of Carnap ([1950b](#) , 67) and Hempel ([1965](#) , 277) are empirical atomic sentences or their negations, for example, this tree is (not) green. Basic sentences in the sense of Popper ([1935](#) /2002, ch. 5, sec. 28) are localized existential sentences, whose matrix is a conjunction of basic formulas, for example, “there is a black raven there.” All basic sentences are observation sentences, but not vice versa. In the Vienna Circle, it was hoped that basic sentences would serve to demarcate the *immediately* observable from that which can only be observed in a *mediated* way, via deduction from that which is immediately observable. For example, disjunctions such as “there is a crow or a raven sitting there” are not immediately observable, but they can be deduced from immediate observation sentences. Closer analysis, however, shows that this demarcation is problematic.

3.3 Logical Truth and Deductive Logic

According to philosophical tradition, logical truths are understood to mean “necessities of thought.” Both the history of philosophy and cognitive psychology show, however, that our *intuitive* ideas about necessities of thought are unreliable. Modern logic replaces the intuitive concept of necessity with exact concepts. The distinction introduced in [sec. 3.1.1](#) between logical and nonlogical concepts is of fundamental importance for the modern definition of logical truth. We will now present a general characterization of logical truth, which is applicable to any system of logic.

([3.3-1](#)) *Logical truth – preliminary characterization*: a sentence is logically true iff its truth follows logically from its syntactic structure and the meaning of its logical concepts alone.

In an equivalent reformulation of this idea, one can say: a sentence is logically true iff its truth follows from its *logical form* (this characterization goes back to [Wittgenstein’s Tractatus of 1921](#)). We find the logical form of a sentence by replacing all of its nonlogical concepts with metalanguage *variables* . The logical form of a sentence is what we get after the *formalization* of the sentence ([sec. 3.1.2](#)), but without any natural language interpretation of its nonlogical terms.

The characterization ([3.3-1](#)) is only preliminary, as the words “follows logically” in the defining phrase make it circular. The essential thing about a logical sentence form is that its nonlogical symbols may be *interpreted* in arbitrary ways. If every such interpretation makes the sentence true, then the sentence is logically true. We can therefore conclusively define the logical truth of a sentence, building on the idea in ([3.3-1](#)), as follows:

(Def. 3.3-1) A sentence is *logically true* (for short: *L-true*) iff its (logical) sentence form is true for every possible interpretation of its nonlogical symbols. We then also call this sentence form logically true.

Definition ([3.3-1](#)) goes back to Tarski ([1936b](#)). It is easily applicable to specific systems of logic, by specifying the permissible interpretations. In PropL, these are the assignments of truth values to atomic sentences. In PredL, these are the interpretations (D,I) explained in [sec. 3.1.2](#) . Note that in possible extensional interpretations (D,I), not only the interpretation function $I(\phi)$ for nonlogical symbols ϕ may vary, but also the individual domain D. Here are some examples:

(3.3-2) *Example of an L-true sentence*:

If everybody is mortal, then nobody is immortal.

Sentence form: if all Fs are Gs, then there exists no F which is not a G.

Formalization: $\forall x(Fx \rightarrow Gx) \rightarrow \neg \exists x(Fx \wedge \neg Gx)$

Further examples: $A \vee \neg A, A \rightarrow (A \vee B), \forall x Fx \rightarrow Fa, Fa \rightarrow \exists x Fx$.

The sentence form ([3.3-2](#)) is L-true, because it is true for every (extensional) interpretation of “F” and “G.” In other words, it is true in every possible world, as extensional interpretations correspond to possible worlds.

In analogy to [def. 3.3-1](#) , one defines a sentence as *logically false* (L-false) iff its sentence form is false for every possible interpretation of its nonlogical symbols. Examples of L-false sentence forms are $p \wedge \neg p$ (e.g. it will rain today and it will not rain), or $p \leftrightarrow \neg p$.

L-false sentences are also called *inconsistent* , or *contradictory* , because one may derive contradictions from them of the form “p and not p.” Sentences which are not L-false are called *consistent* . Consistency is also defined for *sets* of sentences: a sentence set is called consistent iff at least one interpretation verifies all sentences in the set; otherwise, it is called inconsistent (or contradictory). Finally, a sentence which is neither L-true nor L-false is called logically *contingent* . Examples of logically contingent sentences are Fa (e.g. Peter is ill) or $\forall x(Fx \rightarrow Gx)$ (e.g. all ravens are black). Most sentences are contingent.

With the help of the same method, we define the logical *validity* of arguments or inferences, which consists in the fact that the *truth* is transferred from the premises to the conclusion with *certainty* , that is, in all possible interpretations or ‘worlds’:

(Def. 3.3-2) An inference is *logically valid* (for short: valid) iff for the logical *form* of this inference the following holds: every possible interpretation which makes all premises true, also makes the conclusion true. – *Equivalently expressed*: ... iff there is no interpretation which makes all the premises true, but the conclusion false. – If an inference is valid, then we say that its conclusion *follows* (logically) from its premises.

(3.3-3) *Example of a logically valid inference*: Inference form:

Premise 1: All humans are mortal. $\forall x(Hx \rightarrow Mx)$

Premise 2: You are a human. Ha

Conclusion: Therefore, you are mortal. Ma

As with logical truth, logical validity is a matter of *logical form*. This fact implies that the validity of an inference does *not* depend on the *truth* of its premises (a common misunderstanding). The fact that an inference is valid only means: *if* its premises *were true*, then its conclusion *would* also have to be true. An inference with false premises can also be valid, like for example: “All humans are rich. You are a human. Therefore you are rich.” This inference is logically valid, since in all possible worlds in which all humans were rich, you being a human would be rich, too. That in the actual world not all humans are rich is merely a contingent fact. Of course, in order to apply an inference for purposes of *justification* or *prediction*, one must have good reasons for believing that its premises are true. Valid inferences, however, may also be used for purposes of *falsification*: in such cases, one infers from the falsity of the conclusion that *at least one* of the premises must be false.

To demonstrate that an inference form is semantically *invalid*, it suffices (according to [def. 3.3-2](#)) to find an interpretation that makes all the premises true but the conclusion false. On the other hand, semantic proofs of the *validity* in predicate logic are complicated and abstract. For this reason, logic also uses another procedure: the method of *deductive proofs*. With this method, one assumes some fundamental L-truths as *axioms* and some L-valid inferences as *rules* and aims to derive all other L-true sentences or valid inferences through the iterative application of axioms and rules in the form of “proofs” (see, e.g., [Klenk 1989](#), [Copi et al. 2010](#)). A full logical system consists of a language, a logical semantics, and a deductive proof system (or calculus).

3.4 Conventions of Meaning and Definitional Truths

As with L-true sentences, definitions such as “A bachelor is a hitherto unmarried man” are *analytical* truths, that is, their truth follows from the meaning of their concepts alone, independently of the factual constitution of the world. Differently from L-true sentences, however, the truth of definitions is determined not by the meaning of their logical concepts, but by conventions of meaning for their specific *nonlogical* concepts, in our example “bachelor,” “man,” and “hitherto unmarried.” The truth of sentences which are definitionally true is therefore *not* a matter of their logical form, but it is a matter of the conventions of meaning in the given language. So we define:

(Def. 3.4-1) (a) A sentence is *definitionally true* (or: *extralogically analytically true*) iff its truth is based on certain *conventions of meaning* for its *nonlogical* concepts.

In other words: a sentence is definitionally true iff it follows logically from the set of these conventions of meaning, without being itself logically true.

(b) A sentence is *analytically true* iff it is logically or definitionally true.

(3.4-1) *Examples of definitionally true sentences*:

a. All bachelors are male. $\forall x(Bx \rightarrow Mx)$

b. The length of the standard measure in Paris is one meter. $\text{Length}(a) = (1, m)$

(3.4-2) *Examples of synthetically true sentences:*

- a. All polar bears are white. $\forall x(Px \rightarrow Wx)$
- b. The length of this rod is one meter. $\text{Length}(b) = (1, m)$

The difference between analytic and synthetic truths can be illustrated as follows: if a zoologist sees the truth of (3.4-2)(a) as empirically confirmed, then his justification makes sense. But if someone reported to us that in spite of extensive empirical investigation, he has never met a female bachelor, then this person has clearly not understood the meaning of “bachelor”: bachelors are by definition (unmarried) men; no empirical research is necessary to know this. The examples also make it clear that definitional truths cannot be distinguished from synthetic truths on the basis of logical form, as (3.4-1)(a) and (3.4-2)(a) or (3.4-1)(b) and (3.4-2)(b) have the same logical form, respectively.

Although definitions cannot be *empirically* true or false, they can be more or less *empirically adequate* (see also [Hempel 1965](#), [sec. 6.5](#)). Before Linné, for example, animals and plants were classified by their phenomenological characteristics. Classification by common evolutionary descent, founded by Darwin a century later, turned out to be substantially more adequate, because it reflected a much larger number of common nomological characteristics. Nevertheless, nobody can be forbidden to define “fish” as “water-dwelling vertebrates.” The fact that dolphins, etc., would then also count as fish, although they descended from land-dwelling mammals, with whom they share many more common features than with the other (gill-breathing) fishes, does not disprove this definition – it does however, highlight its empirical inadequacy.

Some authors call particularly natural or empirically adequate definitions ‘true’ or ‘real definitions,’ as distinct from “nominal definitions” in the sense of linguistic conventions. This view goes back to Aristotle and presupposes an *essentialism*, according to which the world holds in itself a ‘true’ partition into natural kinds, which are characterized by certain necessary or “essential” properties (cf. [Fine 1994](#); [Schrenk 2017](#), [sec. 6.3](#)). What one counts as “essential” is however often unclear: there is ongoing debate about the optimal classification of biological taxa. Purely evolutionary classifications are generally held as superior to Linnaean classifications, but they would have to classify crocodiles not together with reptiles, but with birds, which is regarded as ‘unnatural’ by most biological taxonomists (cf. [Ridley 1993](#), 369). This shows that optimal definitions are not only dictated by the facts of the world, but also depend on epistemic matters. For this reason, we do not adopt the notion of ‘real definitions,’ because the demarcation between definitions and synthetic sentences would thereby become *unclear*. By “definitions,” we always understand more or less adequate semantic conventions.

There are three different kinds of sentences which are definitionally true, namely:

(D1) *Explicit definitions*: In these, a concept (the definiendum) is equated by convention with a complex of other concepts (the definiens). For general concepts, an explicit definition has the form of an *equivalence* (see [Suppes 1999](#), sec. 8.3):

(3.4-3)

(Definiendum) $\underbrace{\hspace{10em}}$	$\leftrightarrow_{\text{def}}$	(Definiens) $\underbrace{\hspace{10em}}$
Example: $\forall x: x \text{ is a bachelor}$ Logical form: $\forall x(Dx_1 \dots x_n \leftrightarrow A)$, where in A only the individual variables $x_1 \dots x_n$ are free.		$x \text{ is a hitherto unmarried man}$

For singular concepts and function concepts, explicit definitions have the form of an identity, for example, $a =_{\text{def}} f(b_1, \dots, b_n)$, or $\forall x(f(x) =_{\text{def}} a \cdot x^n + b)$.

(D2) *Meaning postulates* (also called ‘*implicit*’ definitions) : These are semantic conventions, which like definitions play the role of extralogical-analytic *axioms*, that is, are not arrived at by logical derivation. These meaning postulates, however, do not have the form of an explicit definition, but some other form, mostly that of an implication. *Example*:

$\forall x(x \text{ is round} \rightarrow x \text{ has no corners})$.

(D3) *Derived definitorially true sentences*: These sentences are not logically true, but follow logically from sentences of type D1 or D2.

Example:

$\forall x(x \text{ is not male} \rightarrow x \text{ is not a bachelor})$.

Proper definitions must fulfill the criterion of *eliminability* of defined concepts (cf. Suppes 1957, 154). In order to satisfy this criterion, they have to satisfy the *noncircularity* requirement. For a single explicit definition, this criterion demands that the definiendum concept must not occur in the definiens. If this is the case, then the definiendum of explicit definitions of the form (3.4-3) can be eliminated in every sentence S in which it occurs, simply by replacing it by its definiens: the result is a sentence S* which is definitionally equivalent with S, that is, the equivalence $S \leftrightarrow S^*$ follows from the given definition.² In distinction to noncircular explicit definitions, meaning postulates do *not* fully characterize a concept and do not allow for its elimination.

For an entire *set* of definitions $\{D_1, \dots, D_n\}$, it must be required that this set can be arranged in the form of a *noncircular chain* of definitions, which means that every definiens in this chain contains only concepts which are either undefined or are the definiendum of an *earlier* definition in the chain (cf. von Savigny 1976, 116). An example of a noncircular chain of definitions is $Ax \leftrightarrow_{\text{def}} Bx \wedge \neg Cx$, $Bx \leftrightarrow_{\text{def}} Dx \wedge Ex$, $Dx \leftrightarrow_{\text{def}} \neg Fx$. If we carry out successive replacement in the definiens of Ax , we arrive at $Ax \leftrightarrow_{\text{def}} \neg Fx \wedge Ex \wedge \neg Cx$ as the basic definition of Ax , which contains only primitive (undefined) concepts in its definiens. An example of a circular chain of definitions is $\text{Necessary}(A) \leftrightarrow_{\text{def}} \neg \text{Possible}(\neg A)$, $\text{Possible}(A) \leftrightarrow_{\text{def}} \neg \text{Necessary}(\neg A)$, as after replacement, we arrive at the circular basic definition $\text{Necessary}(A) \leftrightarrow_{\text{def}} \neg \neg \text{Necessary}(\neg A)$.

A further category of definition is *recursive* definitions. They play an important role in logic and computer science. Without going into details, we only mention that in spite of their iterative structure, recursive definitions also satisfy the noncircularity requirement, since they are provably equivalent with (quite complicated) explicit definitions (cf. van Dalen et al. 1978, 157).

The central requirement for analytically true sentences is their empirical non-creativity, which is explicated as follows:

(Prop. 3.4-1) *Requirement of empirical non-creativity*: Analytically true sentences must neither have empirical content, nor create new empirical content in combination with the given system S of accepted background beliefs.

More precisely: A set of sentences Δ can *only* be analytically true if (a) the empirical content of Δ (in the sense of def. 3.7-1) is empty, and (b) the empirical content of S-united-with- Δ is the same as that of S.

Both requirements (a) and (b) must be fulfilled, in order to be able to say that the truth of such sentences is based on analytical convention. In Suppes' criterion (1999, 154) requirement (a) is missing, because he deals only with non-creativity *relative* to S.

Prop. 3.4-1 implies that *no concept may be defined twice* in different ways. It is well known that in the history of measurement, different conventions were suggested. Here are two examples:

- 1 meter = the length of the Platinum-Iridium-bar in Paris.
2. 1 meter = the length of a pendulum at sea level with an oscillation frequency of one second.

(1) + (2) however, implies the following *empirical consequence* : the length of the Platinum-Iridium-bar in Paris equals the length of a pendulum at sea level with an oscillation frequency of one second. We cannot regard (1) and (2) *simultaneously* as the definition of one meter. If (1) is the accepted definition, then (2) is an empirical hypothesis, and vice versa. Many “definitions” to be found in textbooks, for example, “Water = H_2O ,” or “Child = person less than 12 years old,” do not fulfill the requirement of empirical non-creativity. This is because the concepts being defined (“water,” “child”) have already been analytically introduced to common-sense background knowledge in other ways, for example, by ostensive characterizations. Such “definitions” should more rightly be called *explications* in the sense of Carnap (recall sec. 2.1) rather than

definitions. Also so-called “definitions” of theoretical concepts, for example, “force = mass times acceleration,” generate new empirical content and for this reason are not proper definitions (see [sec. 5.1](#)).

As long as a concept in a system of knowledge is characterized only by *one* explicit definition, this definition is provably always empirically non-creative (see [Shoenfield 1967](#) , 57f). Also a *single* implicational meaning postulate for a new concept Fx such as $\forall x(Fx \rightarrow Ax)$ is not creative, because it is logically weaker than an explicit definition $\forall x(Fx \leftrightarrow Ax)$. Mostly though, an entire *system* of meaning postulates is introduced for one (or several) concepts; and for this, there is no guarantee of non-creativity. The designation of systems of meaning postulates as “implicit definitions” (which goes back to Hilbert) is therefore easily misunderstood, because “implicit definitions” behave quite differently from proper definitions.

3.5 Classification of Sentences According to Their Generality

The overview is shown in [fig. 3.5-1](#) . The logical *strength* decreases successively with the transition from universal via singular to existential sentences: a universal sentence “Everything is an F” is logically stronger than its singular instance “this is an F” which is in turn logically stronger than the existential sentence “something is an F.”

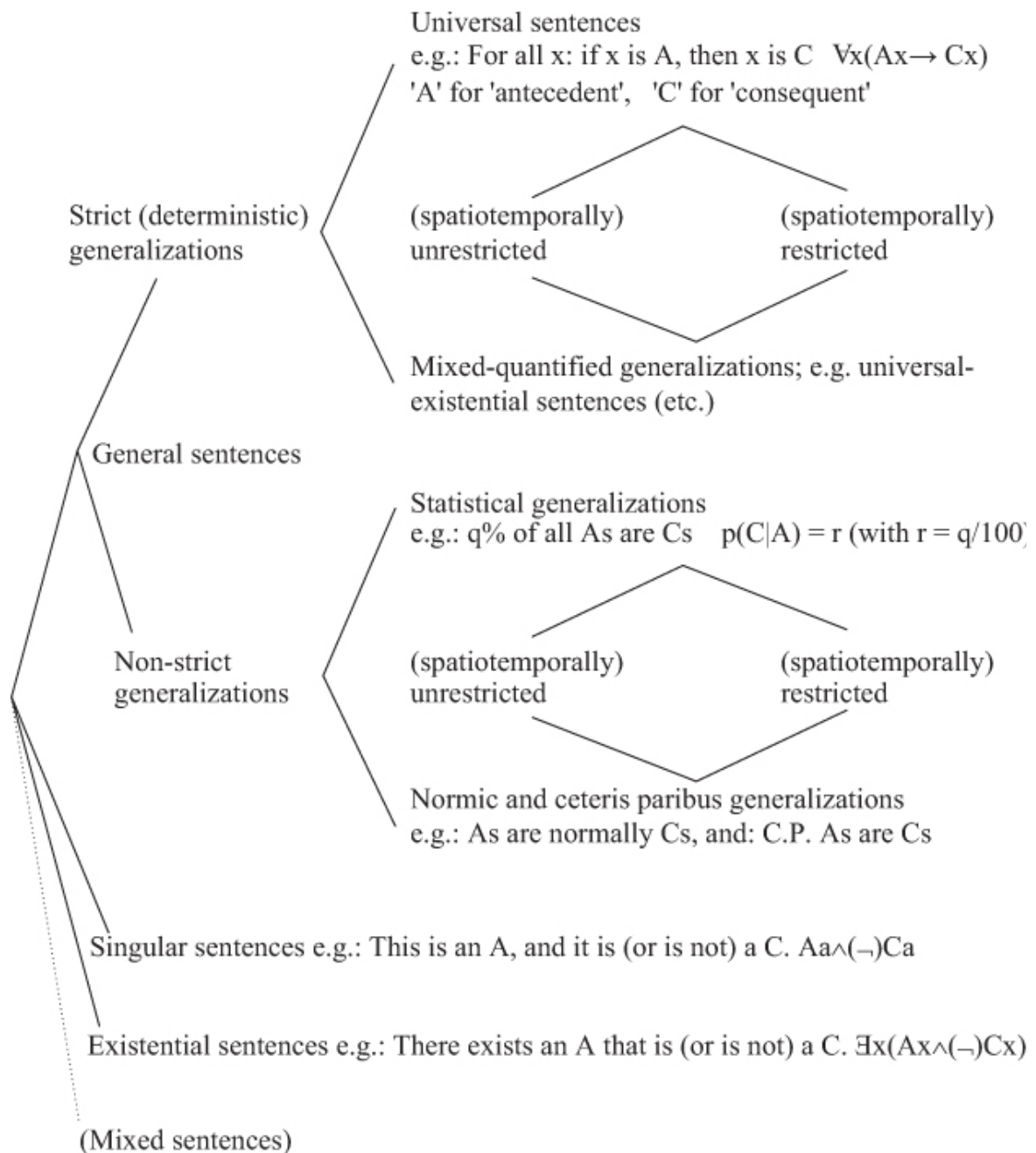


Figure 3.5-1 Classification of sentences according to their generality. ↩

1. *Strict generalizations* contain at least one universal quantifier. They include:

The *universal implication* “For all x: if Ax, then Cx” is the simplest form of a universal sentence. The if-component Ax is also called the *antecedent*, the then-component is called the *consequent*. In general, a *purely* universal sentence begins with one or more universal quantifiers, followed by a quantifier-free formula, which is called the *matrix* of the universal sentence and typically has the form of an implication, an equivalence, or a numerical identity. Universal sentences are also called *strict* or *deterministic* or *exceptionless* generalizations, since the truth of $\forall x(Ax \rightarrow Cx)$ implies for *every* individual a that if Aa obtains, the occurrence of Ca is inevitably determined.

Mixed-quantified generalizations contain both universal and existential quantifiers. *Example*: every action has a motive, formally $\forall x \exists y(\text{Action}(x) \rightarrow \text{Motive}(y, x))$. The meaning of generalizations with mixed

quantifiers depends on the *order* of quantifiers: $\exists x \forall y Rxy$ is a logically stronger statement than $\forall y \exists x Rxy$. *Example:* let “ Rxy ” stand for “ x is the cause of y .” Then $\forall y \exists x Rxy$ says that everything has some cause, but $\exists x \forall y Rxy$ says that some x is the cause of everything, that is, everything has the *same* cause, which is a much stronger statement. If the quantifiers are the same, though, their order is unimportant.

2. *Non-strict generalizations:* In this category, we include primarily *statistical* generalizations of the form “ $q\%$ of all A s are C s.” One calls this a statement of *conditional* probability and writes it down as $p(Cx|Ax) = r$, with $q = r/100$. The probability “ p ” is understood statistically, as a frequency or frequency limit ([sec. 3.9](#)). Although statistical generalizations are general sentences, they are *not* universal sentences. The statistical generalization “ $q\%$ of all A s are C s” says nothing about single individuals of type A , rather it only says something about the class of A s, that is, that $q\%$ of the individuals of the class of A s are located in the class of C s (see [fig. 3.5-2](#)). For this reason, there are no logical consequence relations, but only inductive-probabilistic relations between statistical generalizations and singular sentences, if the domain of individuals is unknown or potentially infinite.

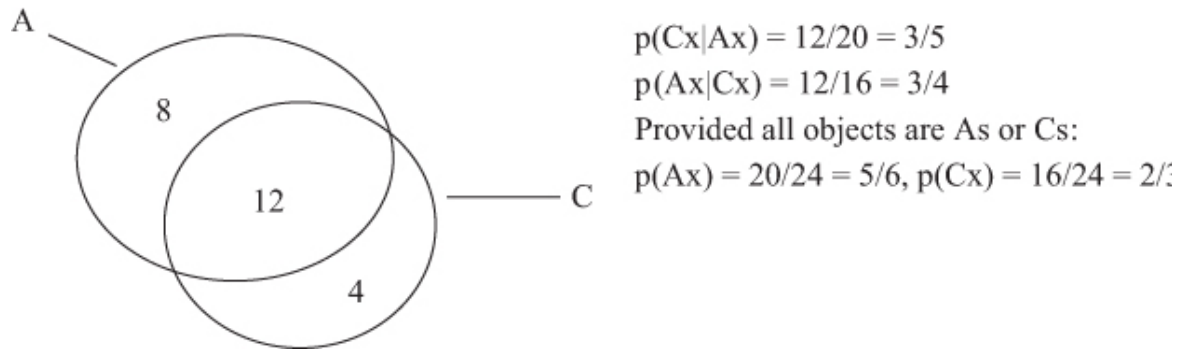


Figure 3.5-2 Statistical generalizations. [↩](#)

A further kind of non-strict generalizations are *normic* generalizations or normal case hypotheses of the form “ A s are *normally* C s” – for example, birds can normally fly (the term “normic” goes back to [Scriven 1959a](#)). For reasons explained in [sec. 4.5.2](#) , we can assume that a normic law analytically *implies* a numerically unspecific statistical generalization of the form “*most* A s are B s,” without being semantically equivalent with it. A third kind of non-strict generalizations are *ceteris paribus* (CP) generalizations, which say that Ax leads to Bx as long as other interfering factors are either *held constant* (comparative CP generalizations), or are *absent* (exclusive CP generalizations, see [sec. 4.5.1](#)).

What statistical, normic, and CP generalizations have in common, is that they allow for *exceptions* – these are instances which verify the antecedent and falsify the consequent predicate. Therefore, inferences from them are described by a non-monotonic consequence relation in the sense of [sec. 2.7.4](#) . Strict laws forbid such exceptions and are falsified by them.

Both strict and non-strict generalizations can be unrestricted or restricted in space and time. They are restricted if the antecedent predicate contains a restriction to a finite area of space and time – for example, “all swans in *Europe* are white,” as opposed to “all swans are white.” Unrestricted universal sentences are also called *spatiotemporally universal* ([sec. 6.5.4](#)).

3. *Singular* are all sentences which contain *no quantifiers* , and also no statistical probability function. Singular sentences are composed of atomic sentences using propositional operators (Fa , $\neg Fa$, $Fa \wedge Rab$, etc.). They always say something about one or more individual things.

4. *Existential sentences:* Pure existential sentences begin with one or more existential quantifiers, followed by a quantifier-free matrix formula. Usually, the matrix has the form of a conjunction $\exists x(Ax \wedge \pm Cx)$; for example, there is a black swan. Such existential sentences may also be unrestricted or restricted, depending on whether Ax contains a restriction in space and time or not.

5. *Mixed sentences* , finally, are propositional combinations of the above sentence types, like for instance $\forall x(Fx \rightarrow Gx) \vee Ha$. They play a subordinate role.

The table in [prop. 3.5-1](#) summarizes the most important logico-deductive relationships between the listed kinds of sentences.

(Prop. 3.5-1) *Notation*: “ $X \Vdash Y$ ” stands for “ Y follows logically from X ” (“ $\dashv$ ” for “logically equivalent”).

1. Universal sentence $\Vdash$ singular sentence
All As are Cs $\Vdash$ if a is A, then a is C
2. Universal sentence and singular sentence $\Vdash$ singular sentence (*explanation scheme*)
All As are Cs, and a is A $\Vdash$ a is C
3. Singular sentence falsifies strictly universal sentence (*falsification scheme I*)
Singular sentence $\Vdash$ negation of a universal sentence
a is A and not C $\Vdash$ not all As are Cs
4. Existential sentence falsifies strictly universal sentence (*falsification scheme II*)
There is an A that is not a C $\Vdash$ not all As are Cs
5. Universal sentence is logically equivalent to the negation of an existential sentence.
All As are Cs $\dashv$ $\dashv$ there is no A that is not a C
6. Singular sentence $\Vdash$ Existential sentence
a is an A $\Vdash$ there is an x that is an A
7. There are no logical consequence relations between non-strict generalizations and singular sentences, only statistical or epistemic relations of probability.

The above characterizations of kinds of sentences are of a *syntactic* nature. Syntactic characterizations are not always invariant with respect to (w.r.t.) logically equivalent (L-equivalent) transformations. For instance, the singular sentence Fa is L-equivalent to the universal sentence $\forall x(x = a \rightarrow Fx)$ (all x which are identical with a have property F). This sentence, however, is not a universal sentence in the *semantic* sense, as it only says something about the individual a . If characterizations of sentences should express genuine properties of the *propositions* thereby expressed, they must be invariant w.r.t. L-equivalent sentence transformation, since L-equivalent sentences express the same proposition. A standard method of solving this problem is the method of *essential* sentence types (see [Hempel/Oppenheim 1948](#) , 271). ³ Here, one regards a sentence as essentially universal (or existential) only if it is not L-equivalent to a singular sentence:

(Def. 3.5-1) *Essential sentence types (according to their generality)*:

1. A sentence S is an essentially universal/existential sentence iff S is a syntactically universal/existential sentence and not L-equivalent to a syntactically singular sentence.
2. By contrast, a sentence S is an *essentially* singular sentence iff S is L-equivalent to a syntactically singular sentence.

The division of kinds of sentences according to their generality is largely *independent* of the content-based classification of sentences given in [sec. 3.2](#) : the kinds of sentences listed above can be either empirical or theoretical sentences, depending on which nonlogical concepts occur in these sentences, and which range their quantifiers possess. Unrestricted universal or existential sentences can however never be observation sentences.

A further distinction, which runs *crosswise* to our distinctions up to now, is the one between *evidences* and *hypotheses* . It relates to the *epistemic status* of our knowledge. Evidence statements are very well-secured sentences – actual observation sentences are prototypical for this category. Hypotheses are less well-secured sentences, which are in need of further testing – prototypical for this are general sentences, or potential observation sentences (i.e. observable predictions).

PHILOSOPHY OF SCIENCE

A Unified Approach

GERHARD SCHURZ

Second Edition



PHILOSOPHY OF SCIENCE

A Unified Approach

GERHARD SCHURZ

Second Edition



“I consider this book to be a unique work both with regard to the sheer amount of material it covers but also with regard to the clarity, depth and insight with which these matters are dealt with by one of Europe’s foremost analytic philosophers. I would not hesitate to use it in an advanced course on the subject, and I believe it would work well also for an introductory course given the careful way in which the book is structured into introductory and advanced material.”

– **Erik J. Olsson** , *Lund University*

“ *Philosophy of Science: A Unified Approach* revamps and bolsters the logical empiricist way to think about and to practice philosophy of science. It combines technical precision and rigor with philosophical breadth and depth. It delves into all major philosophical problems and issues in general philosophy of science. It elegantly presents the received positions and approaches, but it also articulates novel views and arguments. It systematically explains the necessary technical concepts and makes them do deep philosophical work. *Philosophy of Science: A Unified Approach* lives up to its promise to highlight and defend the methodological unity of the sciences. Gerhard Schurz has written a superb introduction to philosophy of science; novices and experts alike have a lot to learn from it.”

Stathis Psillos , *University of Western Ontario and University of Athens*

“This is an excellent introduction to the general philosophy of science. It is lucidly written and very well structured in elementary and advanced subsections. Inspired by the classic authors of the 20th century, such as Carnap and Popper, and guided by the methodological version of the idea of the unity of the sciences and the humanities, Gerhard Schurz shows how it stands with the main topics in the first decades of the 21st century.”

– **Theo Kuipers** , *University of Groningen*

Philosophy of Science

Philosophy of Science: A Unified Approach, Second Edition, combines a general introduction to philosophy of science with an integrated survey of all its important subfields. As this book's subtitle suggests, this excellent overview is guided methodologically by "a unified approach" to philosophy of science. Behind the diversity of scientific fields, one can recognize a methodological unity. This unity is worked out in this book, revealing all the while important differences between subject areas.

This book offers a two-part approach that makes it an excellent introduction for readers new to the field and a useful resource for more advanced students. Each chapter is divided into two sections. The first section assumes no foreknowledge of the subject introduced, and the second section builds upon the first by bringing into the conversation more advanced, complementary topics from contemporary research.

Definitions, key propositions, examples, and figures complement all of the core material. At the end of every chapter, there are selected readings and exercise questions (with solutions and answers to many of them at the end of this book). This book also includes a comprehensive bibliography and an index.

The Second Edition has been revised to make the technical material much more accessible for students outside of the sciences and philosophy. It also incorporates new research over the last decade, including significantly expanded coverage of values in science, mechanistic explanations, and inductive risk. An extended index of subjects that can also be used as a glossary has been added to the back of this book.

Gerhard Schurz (MSc 1980, PhD 1983) is Professor of Philosophy at the Heinrich Heine University Düsseldorf (since March 2002) and member of the Leopoldina (German Academy of Sciences) and the International Academy for Philosophy of Science. Before 2002, he was an Associated Professor at the University of Salzburg and a Guest Professor at the University of California at Irvine and Yale University. His research areas include philosophy of science, epistemology, logic, and cognitive science. He is the author of 270 research articles and 12 books.

Philosophy of Science

A Unified Approach

Second Edition

Gerhard Schurz



Designed cover image: Getty Images

Second edition published 2026
by Routledge
605 Third Avenue, New York, NY 10158

and by Routledge
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2026 Gerhard Schurz

The right of Gerhard Schurz to be identified as author of this work has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

The Open Access version of this book, available at www.taylorfrancis.com, has been made available under a Creative Commons Attribution-ShareAlike 4.0 (CC BY-SA) International license.

Any third party material in this book is not included in the OA Creative Commons license, unless indicated otherwise in a credit line to the material. Please direct any permissions enquiries to the original rightsholder.

For Product Safety Concerns and Information please contact our EU representative
GPSR@taylorandfrancis.com. Taylor & Francis Verlag GmbH, Kaufingerstraße 24, 80331 München, Germany.

Trademark notice : Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

First edition published by Routledge 2013

Library of Congress Cataloging-in-Publication Data

Names: Schurz, Gerhard, 1956– author

Title: Philosophy of science : a unified approach / Gerhard Schurz.

Other titles: Einführung in die Wissenschaftstheorie. English

Description: 2nd edition. | New York, NY ; Abingdon, Oxon : Routledge, 2026. | Includes bibliographical references and index. | Text in English. Translation from German.

Identifiers: LCCN 2026005541 | ISBN 9781041145820 hardback | ISBN 9781041145813 paperback | ISBN 9781003675075 ebook

Subjects: LCSH: Science—Philosophy

Classification: LCC Q175 .S413414 2026

LC record available at <https://lcn.loc.gov/2026005541>

ISBN: 978-1-041-14582-0 (hbk)

ISBN: 978-1-041-14581-3 (pbk)

ISBN: 978-1-003-67507-5 (ebk)

DOI: [10.4324/9781003675075](https://doi.org/10.4324/9781003675075)

Typeset in Sabon
by codeMantra

Contents

[List of Figures, Definitions, and Key Propositions](#)

[List of Symbols](#)

[Foreword](#)

[Foreword to the Second Edition](#)

[1 Introduction: Where Do We Stand?](#)

[1.1 Tasks and Aims of Philosophy of Science](#)

[1.2 On the Historical Development of Philosophy of Science](#)

[1.3 Philosophical Positions in Contemporary Philosophy of Science](#)

[1.3.1 Logical Empiricism](#)

[1.3.2 Critical Rationalism](#)

[1.3.3 The Historical Account and Relativism](#)

[1.3.4 Metaphysical and Scientific Realism](#)

[1.3.5 Naturalism and the Cognitive Turn](#)

[1.3.6 Structuralist and Semantic View of Theories](#)

[1.3.7 Pragmatic Philosophy of Science](#)

[1.3.8 Hermeneutics and Critical Theory](#)

[1.3.9 Radical Constructivism, Postmodernism, and the Science Wars](#)

[1.4 Introductory Literature and Exercises](#)

[2 The Question of Unity: In Search for Common Foundations of the Sciences](#)

[2.1 Normative or Descriptive? The Method of Rational Reconstruction](#)

[2.2 Common Epistemological Assumptions](#)

[2.3 Common Methodological Features](#)

[2.4 Scientific Disciplines and Their Classification](#)

[2.4.1 The Special Status of Mathematical Sciences](#)

[2.4.2 Testing the Common Foundations of Science](#)

[2.4.3 Two Further Classifications of Sciences](#)

[2.5 The Question of Value Neutrality](#)

[2.5.1 Max Weber's Value-Freedom Postulate](#)

[2.5.2 Value Neutrality and Means-End Inferences](#)

[2.5.3 Explication of the Value-Neutrality Requirement](#)

[2.6 The Demarcation Problem and the Methodological Unity of Sciences](#)

[2.6.1 Demarcation](#)

2.6.2 Methodological Unity instead of Physical Reductionism

2.7 Scientific Inference and Argumentation

2.7.1 Deduction and Induction

2.7.2 Popper and the Significance of Inductive Inference in the Sciences

2.7.3 Abduction and Inference to the Best Explanation

2.7.4 Monotonic and Non-monotonic Inferences

Complementary and Advanced Topics

2.8 How Much Metaphysics is Presupposed by Science?

2.8.1 Truth, Idealism, and Realism

2.8.2 Constructive Realism versus Radical Constructivism

2.9 On the Theory (In)dependence of Observation

2.9.1 Theory-Dependency of Observation – Pros and Cons

2.9.2 A Definition of Theory-Neutral Observation Concepts

2.10 Value Neutrality in the Context of Contemporary Debates

2.10.1 The Role of Values in Science

2.10.2 The Argument of Inductive Risk

2.10.3 Meta-ethical Justification of Value Neutrality and the Difference between Experience and Valuation

2.11 The Justification of Induction: An Unsolvable Problem?

2.11.1 Is Induction Rationally Unjustifiable?

2.11.2 The Optimality of Meta-induction

2.12 Introductory Literature and Exercises

3 The Conceptual Toolkit: Language, Logic, and Probability

3.1 Kinds and Classification of Concepts

3.1.1 The Logical Type of Concepts

3.1.2 Syntax and Semantics, Statements, and Models

3.1.3 The Content Type of Concepts

3.1.4 The Gradation (Scale) Type of Concepts

3.2 Classification of Sentences According to Their Content Type

3.2.1 Descriptive versus Prescriptive Sentences

3.2.2 Observation and Basic Sentences, Empirical and Theoretical Sentences

3.3 Logical Truth and Deductive Logic

3.4 Conventions of Meaning and Definitional Truths

3.5 Classification of Sentences According to Their Generality

3.6 General Sentences, Lawlikeness, Determinism, and Indeterminism

3.7 Content of Sentences and Kinds of Content

3.8 Verification, Falsification, Confirmation, and Disconfirmation

3.9 Objective (Statistical) and Subjective (Epistemic) Probability

Complementary and Advanced Topics

3.10 Disposition Concepts

3.11 Challenges to the Logical-Definitional-Synthetic Distinction

- [3.12 Relevance and Irrelevance in Logical Inferences](#)
- [3.13 More on Probability](#)
 - [3.13.1 Mathematical Laws of Probability](#)
 - [3.13.2 Problems of Objective-Statistical Probability Concepts](#)
 - [3.13.3 Problems of Subjective Probability Concepts](#)
 - [3.13.4 Connections between Objective and Subjective Probabilities: Principal Principle and Narrowest Reference Class](#)
- [3.14 Introductory Literature and Exercises](#)

4 A Question of Fit: Law Hypotheses and Their Empirical Testing

- [4.1 The Condition of Relevance](#)
 - [4.1.1 Relevance of Strict Laws](#)
 - [4.1.2 Relevance of Statistical Laws](#)
- [4.2 Testing of Strict Law Hypotheses for Truth and Relevance](#)
 - [4.2.1 The Method of Agreement and Difference](#)
 - [4.2.2 Methodological Induction: How to Discover Strict Law Hypotheses](#)
- [4.3 Testing Statistical Laws](#)
 - [4.3.1 Testing for Likely Truth and Acceptance Intervals](#)
 - [4.3.2 Discovery of Statistical Hypotheses and Confidence Intervals](#)
 - [4.3.3 Test Statistics and Inference Statistics](#)
 - [4.3.4 Testing for Relevance and Significant Differences](#)
 - [4.3.5 Statistical Representativeness](#)
 - [4.3.6 Sources of Error in Statistical Methods](#)
 - [4.3.7 Applying Statistical Hypotheses to Individual Cases](#)
- [4.4 Correlation and Causality](#)
 - [4.4.1 Hidden Variables](#)
 - [4.4.2 Causal Direction](#)

Complementary and Advanced Topics

- [4.5 Ceteris Paribus Laws and Normic Laws](#)
 - [4.5.1 Comparative versus Exclusive CP Laws](#)
 - [4.5.2 Normic Laws and Their Evolution–Theoretic Foundation](#)
- [4.6 Probability Distributions for Continuous Variables](#)
- [4.7 From Classical Statistics to Bayesian Statistics](#)
 - [4.7.1 The Likelihood Intuition and Its Bayesianism Justification](#)
 - [4.7.2 Objective Bayesianism and the Principle of Indifference](#)
 - [4.7.3 Subjective Bayesianism and Convergence of Opinions](#)
- [4.8 Introductory Literature and Exercises](#)

5 Going beyond Experience: Theories and Their Empirical Evaluation

- [5.1 Theoretical Concepts and Multiple Laws of Correspondence](#)
- [5.2 The Example of Newtonian Physics](#)
- [5.3 The Structure of Scientific Theories](#)
 - [5.3.1 Components of Theories](#)
 - [5.3.2 Theory Nets, Intertheoretical Relations, and Kinds of Theoretical Concepts](#)

- 5.4 [Methodological Features of \(Good\) Scientific Theories](#)
 - 5.4.1 [System Character: Holism of Meaning, Content, and Theory Testing \(Duhem's Thesis\)](#)
 - 5.4.2 [Empirical Creativity, Global Unification, and Novel Predictions: Answers to Hempel's Dilemma](#)
- 5.5 [The Example of Piaget's Theory of Cognitive Development](#)
- 5.6 [The Dynamics of Scientific Theories](#)
 - 5.6.1 [Lakatos' Model of Theory Revision](#)
 - 5.6.2 [Theory Evaluation and Theory Progress](#)
- 5.7 [Instrumentalism and Realism: The Ontology of Scientific Theories](#)

[Complementary and Advanced Topics](#)

- 5.8 [Instrumentalism-Realism Controversies in Philosophy of Science](#)
 - 5.8.1 [The Ramsey Sentence of a Theory: An Instrumentalist Challenge](#)
 - 5.8.2 [The Realist Interpretation of the Ramsey Sentence](#)
 - 5.8.3 [The No-Miracles Argument](#)
 - 5.8.4 [The Pessimistic Meta-Induction](#)
 - 5.8.5 [The Empirical Underdetermination of Theories](#)
 - 5.8.6 [Intertheoretical Correspondence, Structural and Selective Realism](#)
 - 5.8.7 [Criteria for Empirical Significance](#)
 - 5.8.8 [Probabilistic Analysis of the NMA, Alternative Theories, and Local Realism](#)
- 5.9 [General Accounts of Confirmation](#)
 - 5.9.1 [Hypothetico-Deductive Confirmation](#)
 - 5.9.2 [Probabilistic Confirmation and Bayesianism](#)
 - 5.9.3 [Bayesian Pseudo-Confirmation through Content Cutting versus Genuine Confirmation](#)
 - 5.9.4 [Goodman's Paradox](#)
 - 5.9.5 [Confirmation and Acceptance](#)
- 5.10 [Testing Theories: Parameter Fitting and Use-Novel Evidence](#)
 - 5.10.1 [Post-Factum Speculations](#)
 - 5.10.2 [Curve Fitting](#)
- 5.11 [Non-Confirmational Accounts of Theory Evaluation](#)
 - 5.11.1 [Truthlikeness](#)
 - 5.11.2 [Unification, Coherence, and Simplicity](#)
- 5.12 [The Problem of Language Dependence](#)
- 5.13 [Introductory Literature and Exercises](#)

[6 In Search of Causes: Explanation and All That Goes with It](#)

- 6.1 [The Deductive-Nomological Model of Explanation](#)
- 6.2 [Explanation versus Prediction and Justification](#)
 - 6.2.1 [Predictions without Explanatory Function](#)
 - 6.2.2 [Causality and Lawlikeness](#)
 - 6.2.3 [Explanations without Predictive Function](#)
 - 6.2.4 [Logical Problems of Irrelevance and Redundancy](#)

6.3 Probabilistic Explanation Models

6.3.1 Inductive-Statistical Explanation

6.3.2 Maximal Specificity

6.3.3 Probabilistic and Causal Relevance Models

6.3.4 Conflicting Intuitions about the Magnitude of Explanatory Probabilities

6.4 Further Types of Explanation

6.4.1 Mechanistic Explanations

6.4.2 Normic Explanations and the Explanation of Human Actions

6.4.3 Explanation and Understanding as Unification

Complementary and Advanced Topics

6.5 Lawlikeness

6.5.1 Laws of Nature versus System Laws

6.5.2 Lawlikeness in the Wide versus Narrow Sense

6.5.3 Lawlikeness i.w.s. and Counterfactuals

6.5.4 Lawlikeness i.n.s., Spatiotemporal Universality, and Maxwell's Condition

6.5.5 The Best System Account

6.5.6 Physical Necessity and Independent Possibility Knowledge

6.6 Causality

6.6.1 Singular and General Causality

6.6.2 Regularity Accounts

6.6.3 Counterfactual Accounts

6.6.4 Causation and the Direction of Time

6.6.5 Causal Process Theories

6.6.6 Interventionist Accounts

6.6.7 Causality as a Theoretical Concept

6.6.8 The Theory of Causal Nets

6.6.9 Causal Discovery and Empirical Content of TCN

6.6.10 From Variables to Events: Overdetermination and Causal Preemption

6.7 Introductory Literature and Exercises

Bibliography

Solutions to Selected Exercises

Index of Names

Index of Subjects

List of Figures, Definitions, and Key Propositions

Figures

- [2.1-1 The method of rational reconstruction](#)
- [2.3-1 The three levels of the scientific method](#)
- [2.4-1 Classification of the factual sciences](#)
- [2.5-1 Schematic representation of the value neutrality requirement](#)
- [2.7-1 The inductive-deductive scheme](#)
- [2.7-2 Interaction of epistemic induction and abduction](#)
- [2.9-1 Duck-rabbit](#)
- [2.9-2 Mueller-Lyer illusion](#)
- [2.9-3 Expected learning curves in ostensive learning experiments](#)
- [3.1-1 Division of concepts according to their logical type](#)
- [3.1-2 Classification of concepts according to their content type](#)
- [3.1-3 Classification of concepts according to their gradation type](#)
- [3.1-4 Ordinal scale](#)
- [3.2-1 Classification of sentences according to their content type](#)
- [3.5-1 Classification of sentences according to their generality](#)
- [3.5-2 Statistical generalizations](#)
- [3.8-1 Testability of generalizations](#)
- [3.9-1 Non-monotonicity of conditional probabilities](#)
- [3.13-1 Three binomial distributions](#)
- [4.1-1 Irrelevance and joke](#)
- [4.3-1 Acceptance interval](#)
- [4.3-2 Acceptance and confidence interval](#)
- [4.3-3 Probability distribution of sample differences and significant difference](#)
- [4.4-1 Common cause](#)
- [4.4-2 Common cause vs. intermediate cause](#)
- [4.4-3 Spurious causal independence](#)
- [4.4-4 What is the direction of causality?](#)
- [4.6-1 Normal distribution](#)
- [4.7-1 Shortest interval](#)
- [4.7-2 Transformation of uniform into non-uniform density](#)
- [5.1-1 One theoretical feature causes many empirical dispositions](#)
- [5.1-2 Beam balances and spring scales on Earth and Uranus](#)
- [5.2-1 Theory lattice in classical mechanics](#)

[5.3-1,2,3 Subdivisions of a theory](#)
[5.3-4 Theory net of physical and chemical theories](#)
[5.6-1 Historical theory sequence](#)
[5.6-2 Theory comparison](#)
[5.6-3 Empirically complementary theories](#)
[5.6-4 Theories of optics after Bunge](#)
[5.10-1 Curve fitting and use-novel evidence](#)
[6.6-1 Explanation of screening-off](#)
[6.6-2 Three statistically strongly indistinguishable graphs](#)
[6.6-3 Causal graph of variables and related graphs of values](#)

Definitions

[\(Def. 2.1-1\) Method of rational reconstruction](#)
[\(Def. 2.5-1\) The general scheme of means-end inferences](#)
[\(Def. 2.7-1\) Monotonicity](#)
[\(Def. 2.9-1\) Theory-neutral observation concepts](#)
[\(Def. 3.2-1\) Descriptive versus prescriptive sentences](#)
[\(Def. 3.2-2\) Observational, empirical, theoretical and T-theoretical sentences](#)
[\(Def. 3.3-1\) Logical truth](#)
[\(Def. 3.3-2\) Logical validity](#)
[\(Def. 3.4-1\) Definitional truth](#)
[\(Def. 3.5-1\) Essential sentence types](#)
[\(Def. 3.7-1\) Logical and empirical content](#)
[\(Def. 3.8-1\) Verification, falsification, confirmation and disconfirmation](#)
[\(Def. 3.9-1\) Statistical and subjective probability](#)
[\(Def. 3.9-2\) Conditional probability](#)
[\(Def. 3.9-3\) Principle of narrowest reference class](#)
[\(Def. 3.10-1\) Bilateral reduction sentence](#)
[\(Def. 3.10-2\) Nomological definition of dispositions](#)
[\(Def. 3.12-1\) Relevance in logical inferences](#)
[\(Def. 3.12-2\) Content elements](#)
[\(Def. 3.12-3\) Relevant content](#)
[\(Def. 3.12-4\) Irreducible representation](#)
[\(Def. 4.1-1\) Condition of relevance for strict law hypotheses](#)
[\(Def. 4.1-2\) Statistical relevance for conjunctive antecedents](#)
[\(Def. 4.3-1\) Acceptance interval](#)
[\(Def. 4.3-2\) Significant difference](#)
[\(Def. 4.6-1\) Mean, variance and standard deviation](#)
[\(Def. 5.3-1\) Kinds of theoretical concepts](#)
[\(Def. 5.8-1\) Ramsey sentence](#)
[\(Def. 5.8-2\) Carnap sentence](#)
[\(Def. 5.8-3\) Ramsey-Lewis sentence](#)
[\(Def. 5.9-1\) Classical HD-confirmation](#)
[\(Def. 5.9-2\) Relevant HD-confirmation](#)
[\(Def. 5.9-3\) Incremental and absolute Bayesian confirmation](#)
[\(Def. 5.9-4\) Genuine confirmation \(full and partial\)](#)
[\(Def. 5.9-5\) Goodman's grue predicate](#)
[\(Def. 5.11-1\) Popper-truthlikeness](#)
[\(Def. 5.11-2\) Schurz/Weingartner-truthlikeness](#)

[\(Def. 6.5-1\) Spatiotemporal universality](#)
[\(Def. 6.6-2\) Counterfactual Causality according to Lewis](#)
[\(Def. 6.6-3\) Interventionist causality after Woodward](#)
[\(Def. 6.6-4\) Screening-off between variables](#)
[\(Def. 6.6-5\) Linking-up between variables](#)
[\(Def. 6.6-6\) Causal Markov axiom](#)
[\(Def. 6.6-7\) Faithfulness condition](#)
[\(Def. 6.6-8\) Minimality and productivity](#)
[\(Def. 6.6-9\) Cause-effect relation between values of variables](#)

Key Propositions

[\(Prop. 2.1-1\) The supreme epistemic goal of science](#)
[\(Prop. 2.5-1\) Value neutrality requirement](#)
[\(Prop. 2.7-1\) General pattern of abduction](#)
[\(Prop. 2.7-2\) Inference to the best explanation](#)
[\(Prop. 2.7-3\) Application rule for monotonic inferences](#)
[\(Prop. 2.7-4\) Application rule for non-monotonic inferences](#)
[\(Prop. 2.11-1\) Universal optimality of AW](#)
[\(Prop. 3.1-1\) Translation between logical and set-theoretical language](#)
[\(Prop. 3.1-2\) Linguistic and model-theoretic theory representation](#)
[\(Prop. 3.4-1\) Requirement of empirical non-creativity](#)
[\(Prop. 3.5-1\) Important logical relationships](#)
[\(Prop. 3.9-1\) Basic axioms of probability](#)
[\(Prop. 3.13-1\) Theorems of unconditional probability](#)
[\(Prop. 3.13-2\) Theorems of conditional probability](#)
[\(Prop. 3.13-3\) Binomial formula](#)
[\(Prop. 3.13-4\) Subjective probabilities as coherent fair betting quotients](#)
[\(Prop. 3.13-5\) Statistical principal principle](#)
[\(Prop. 3.13-6\) Weak induction principles for subjective probabilities](#)
[\(Prop. 4.1-1\) Statistical relevance and correlation for qualitative properties](#)
[\(Prop. 4.2-1\) Testing a strict law hypothesis](#)
[\(Prop. 4.3-1\) Testing a statistical law hypothesis](#)
[\(Prop. 4.4-1\) Reichenbach conditions for screening off](#)
[\(Prop. 4.5-1\) Comparative CP law](#)
[\(Prop. 4.5-2\) Exclusive CP assertion](#)
[\(Prop. 4.5-3\) Normic interpretation of eCP laws](#)
[\(Prop. 4.6-1\) Mean value, variance and standard deviation of the distribution of sample means](#)
[\(Prop. 4.7-1\) Principle of Bayes-statistics](#)
[\(Prop. 4.7-2\) Bayesian justification of the likelihood-intuition](#)
[\(Prop. 4.7-3\) Consequences of the indifference principle](#)
[\(Prop. 4.7-4\) Bayesian convergence results](#)
[\(Prop. 4.7-5\) Biased priors](#)
[\(Prop. 5.1-1\) Synthetic nature of laws of correspondence for theoretical concepts](#)
[\(Prop. 5.1-2,3\) Biased indicators](#)
[\(Prop. 5.4-1\) Logical background of the holism of testing](#)
[\(Prop. 5.4-2\) Empirical creativity](#)
[\(Prop. 5.6-1\) Falsification of theory version](#)
[\(Prop. 5.6-2,3\) Successes and failures of theory versions](#)
[\(Prop. 5.6-4\) Confirmation of theory versions](#)

[\(Prop. 5.6-5\) Dynamical truthlikeness](#)
[\(Prop. 5.6-6\) Theory comparison and theory progress](#)
[\(Prop. 5.6-7\) Rational acceptability of theory versions](#)
[\(Prop. 5.6-8\) Untested empirical excess content](#)
[\(Prop. 5.6-9\) Confirmation of theory cores](#)
[\(Prop. 5.6-10\) Preference between theory cores](#)
[\(Prop. 5.6-11\) Rational acceptability of theory cores](#)
[\(Prop. 5.8-1\) Newman's problem](#)
[\(Prop. 5.8-2\) Howson's analysis of the NMA](#)
[\(Prop. 5.9-1\) HD-confirmation and Bayesian confirmation](#)
[\(Prop. 5.9-2\) Impossibility results for rational acceptance](#)
[\(Prop. 5.10-1\) Necessary criteria for transfer of probability increase](#)
[\(Prop. 6.1-1\) DN-explanation of events](#)
[\(Prop. 6.1-2\) DN-explanation of laws](#)
[\(Prop. 6.2-1\) Lawlikeness and causality in DN-explanations](#)
[\(Prop. 6.2-2\) Predictive function](#)
[\(Prop. 6.2-3\) Logical relevance conditions for DN-explanations](#)
[\(Prop. 6.3-1\) Inductive-statistical explanation](#)
[\(Prop. 6.3-2\) Maximal specificity](#)
[\(Prop. 6.4-1\) Normic explanation](#)
[\(Prop. 6.4-2\) Unification condition](#)
[\(Prop. 6.5-1\) Physical necessity](#)
[\(Prop. 6.6-1\) Singular and general causality](#)
[\(Prop. 6.6-2\) Two versions of the causal Markov axiom](#)
[\(Prop. 6.6-3\) Equivalence of productivity and minimality](#)
[\(Prop. 6.6-4\) Causal discovery algorithm](#)
[\(Prop. 6.6-5\) Empirical content of TCN](#)
[\(Prop. 6.6-6\) Causal nets with time-indexed variables](#)

List of Symbols

$\neg$	not
$\wedge$	and
$\vee$	or
$\rightarrow$	material implication (if-then)
$\leftrightarrow$	equivalent (if-and-only-if)
$\leftrightarrow_{\text{def}}$	equivalent by definition
$\forall x$	for all x
$\exists x$	there exists an x
$\exists!x$	there exists exactly one x
$=$	identical
$\neq$	not identical
$=_{\text{def}}$	identical by definition
$\square$	necessary
$\mathbb{N}$	set of natural numbers
$\mathbb{R}$	set of real numbers
$\in$	is an element of

$\{a_1, \dots, a_n\}$	set of individuals
$(a_1, \dots, a_n)$	ordered sequence of individuals
$\{x: Fx\}$	set of all individuals which possess property F
$\subseteq$	proper or improper subset
$\subset$	proper subset
$\cap$	intersection
$\cup$	union
$-$	applied to sets set-theoretical complement
$\times$	Cartesian product between sets
$\cdot$	ordinary product between numbers
$/$	division (example $x \cdot y / z = x$ times y divided through z)
$\square$	empty set
$\sum_i \in_I x_i$ and $\prod_i \in_I x_i$	sum and product (respectively) over all x_i with i in I
D	domain
D^n	n -fold Cartesian product of D
I	interpretation function
$\mathcal{L}$	formal language
$M(\mathcal{L})$	set of models for language $\mathcal{L}$
$\leq$	smaller or equal
$\geq$	greater or equal
$\equiv$	equivalence relation

$f:A \rightarrow B$	function from domain A into set of values B
$\models$	verifies (makes true)
$\Vdash$	follows logically from
$\nVdash$	does not follow logically from
$\dashv \vdash$	is logically equivalent with
p	statistical probability
P	epistemic probability
$C(S)$ and $C_r(S)$	logical content and relevant logical content of a sentence set S
$E(S)$ and $E_r(S)$	empirical content and relevant empirical content of a sentence set S
Ω	possibility space
AL	algebra
$\ A\ $	semantic extension of formula A
$df(x)/dx$	first derivative of f(x)
$\sqrt{}$	square root
$\int$	integral
$\lim_{n \rightarrow \infty}$	limit for n going to infinity
$R(T)$	Ramsey sentence of theory T
$(IN) \text{ DEP } (X,Y)$	variables X and Y are probabilistically (in)dependent

Foreword

I have written *Philosophy of Science: A Unified Approach* for beginners as well as for graduate students and experts. This twofold orientation is reflected in the major structure of the chapters, which start with introductory sections for beginners, followed by sections entitled “complementary and advanced topics,” in which these topics are explained in a condensed way, including presentations of my own accounts. In this way, this book attempts to offer not only an introduction, but also a detailed overview of the various areas of the Philosophy of Science and their interrelations, which will help the reader familiarize him- or herself with the present state of the art in this field.

This book addresses itself both to philosophers and to scientists of all disciplines. By way of bridging Philosophy of Science with neighboring fields such as statistics, hermeneutics, cognitive science, the debate on science and values, and detailed examples of contemporary theories in sciences and humanities, this book tries to convey a deep understanding of the sciences and their methods. This book is *self-contained* and written without presuppositions about the readers. On rare occasions, formulas are used, but I’ve always explained them in a clear and easy-to-follow language.

Although this book introduces the reader to all major areas of philosophy of science, it also has a bias and a general philosophical message. Concerning the bias, the author understands himself as a follower of logically and empirically oriented philosophers of science such as Carnap, Hempel, Reichenbach, and Stegmüller in the German-speaking area. This book’s general message is reflected in its subtitle, *A Unified Approach*. Behind the diversity of phenomena investigated in disciplines as different as elementary particle physics and social sciences, or chemistry and linguistics, one can recognize a deep common ground that unites all of the sciences. This unity, however, is not so much a unity at the level of objects (as it was claimed by the proponents of physical reductionism), but rather a *methodological unity*. [Chapter 2](#) is concentrated on finding the common methodological features of the sciences and their epistemological presuppositions. The remaining chapters will not only work out these common features in all details but will also reveal some important differences between characteristic groups of disciplines, such as the physical sciences versus special sciences (e.g. biology, psychology, or social sciences).

Use in teaching: By restricting the teaching materials to Chapters 1–5 and omitting all sections on advanced topics, one obtains the basis for a semester-long two-hours-per-week course. Covering all of the material contained in this book will fill two such courses. The *core materials* of this book are highlighted in the form of *definitions* and *key propositions*.

Conventions on numberings: The numbering of figures follows the following pattern: “Fig. SecNo.-No.,” where “SecNo.” is the two-digit section number, and “No.” is the number of the figure within the section. Example: “2.1-2” is the second figure of [sec. 2.1](#). In the same way, *definitions* are numbered with “(Def. SecNo.-No.),” *key propositions* with “(Prop. SecNo.-No.),” and *examples* or other *highlighted* passages with “(SecNo.-No.).” Example: “(Def. 4.2-2)” is the second definition of [sec. 4.2](#); “(3.1-4)” is the fourth example of [sec. 3.1](#). *Quotations:* Double quotes are used for the purpose of literal quotation and single quotes for stylistic quotation (scare quotes).

At the end of every chapter, the reader finds hints on selected literature and exercises. This book concludes with a comprehensive bibliography and several indices.

This textbook is a revised and significantly extended English elaboration of my German Introduction to Philosophy of Science (“Einführung in die Wissenschaftstheorie”). The introduction has been developed from a series of lectures in the Philosophy of Science that I have given at the Universities of Salzburg, Erfurt, Düsseldorf and, less frequently, at other universities (including California at Irvine, Yale, Vienna, Helsinki, Graz, and Fribourg).

Parts of the text of the German edition have been pretranslated into English by Fergal Treanor. For valuable help concerning matters of content, I am indebted to Ioannis Votsis, Paul Thorn, Hannes Leitgeb, Wolfgang Spohn, Martin Carrier, Franz Huber, Markus Werning, Oliver Scholz, Andreas Hüttemann, Markus Schrenk, Hans Rott, Wolfgang Balzer, Gernot Kleiter, Theo Kuipers, Ilkka Niiniluoto, Erik Olsson, Stephan Hartmann, Jon Williamson, Igor Douven, Ken Gemes, James Ladyman, John Worrall, Stathis Psillos, James Hawthorne, Alan Musgrave, Mathias Frisch, Clark Glymour, Kevin Kelly, Horacio Arló-Costa, and Alvin Goldman. What remains on my side is to wish the readers of this book that it shall bring them much fun and intellectual profit.

Gerhard Schurz
Düsseldorf, March 2013

Foreword to the Second Edition

Since 2013, much has happened in the philosophy of science. Therefore, the Second Edition contains a lot of new content, concerning controversies about the role of values in science (sec. 2.10), results about meta-induction (sec. 2.11), debates concerning the question of scientific realism (sec. 5.8), and the mechanistic philosophy of science (sec. 6.4.1). In compensation, other and more technical parts have been shortened. Moreover, new literature has been added all over the place. The exercises with selected solutions have been extended. The index of subjects has been expanded so that it can be used as a *glossary* : for important terms, the page numbers on which the term is defined are shown in **bold**.

I would like to extend my thanks for valuable content-related help from the first edition to include Karel Lambert, Brian Skyrms, Branden Fitelson, Gila Sher, Johan von Benthem, Lorenzo Magnani, Holger Andreas, Mario Günther, Peter Brössel, Tom Sterkenburg, Leah Henderson, Richard Dawid, Matias Osta Vélez, Jochen Briesen, Gustavo Cevolani, Atocha Aliseda, Petr Spelda, Elke Brendel, Corina Strößner, Alexander Gebharder, Christian Feldbacher-Escamilla, Gottfried Vosgerau, Alexander Christian, Karim Baraghith, Niklas Parwez, Avril Styrman, and Christoph Schamberger.

Gerhard Schurz
Düsseldorf, September 2025

1 Introduction

Where Do We Stand?

DOI: [10.4324/9781003675075-1](https://doi.org/10.4324/9781003675075-1)

This chapter has been made available under a CC BY-SA license.

1.1 Tasks and Aims of Philosophy of Science

Science plays a determining role in the life of today's society. In what sense and to what degree can we trust the results of science? There is a scientific discipline, which examines such questions systematically: the *Philosophy of Science*, which in the German language is called "Wissenschaftstheorie" ("Theory of Science"). Following the German usage of the word, I always understand "science" ("Wissenschaft") in its *broad* meaning – including besides natural sciences, also social sciences, humanities, etc.

The discipline of philosophy (or theory) of science investigates how scientific knowledge works – its goals and its methods, its achievements and its limitations. Like any other scientific discipline, philosophy of science has historically developed out of philosophy. In the present days, it is practiced both by philosophers and by scientists of various disciplines.

One distinguishes between the *general* philosophy of science and philosophies of *particular* sciences. The latter ones are concerned with special kinds of disciplines such as philosophy of physics, biology, psychology, the social sciences, or the humanities. The general philosophy of science discovers those components of knowledge, which are more or less common to all scientific disciplines. Its main questions are:

- i. How is a scientific language constructed?
- ii. What are the rules of correct scientific reasoning or argument?
- iii. What are the distinctive features of a scientific observation or 'datum'?
- iv. What is a scientific law, and what does a scientific theory consist of?
- v. How are law hypotheses and theories empirically tested? When do they count as confirmed, and when as refuted?
- vi. What are the criteria for scientific progress?
- vii. What is achieved by a scientific explanation and what is causality?

This book focuses on these questions of the general philosophy of science, and the following chapters are arranged in corresponding order to the above questions. At the same time, we shall also deal with applications and topics of the philosophies of particular sciences, and examine characteristic *differences* between different branches of science. The most *general* questions of philosophy of science include:

- viii Is there an objectively discoverable reality and an objective truth?
- ix What connection exists between science and value judgments?

Question viii lies on the border between philosophy of science and epistemology (see [sec. 2.8](#)), and question ix on the border with meta-ethics (see [sec. 2.5](#), [2.10](#)).

Beyond its fundamental significance, which consists in providing answers to the above questions, philosophy of science also has a number of significant *applications*, both inside and outside science.

The *internal* applications of philosophy of science consist of the supply of foundational knowledge, which aids the scientist's decisions concerning new or controversial questions. Internal applications also consist of working out *interdisciplinary* common ground and providing *argumentative* competence in critical reflection. Not least, philosophy of science has often paved the way for new scientific disciplines (such as probability theory and statistics, measurement theory, or parts of cognitive science).

Among the *external* applications of philosophy of science to problems in a societal context, two in particular should be emphasized:

- a. The *demarcation* problem is highly significant in society. In this context, it consists of the question of which of our ideas have a claim to the status objective scientific knowledge that should be taught in public educational institutions, as opposed to subjective opinion, political values, ideologies, or religious convictions. This question became politically explosive in the controversy surrounding the creationism movement in the United States. For example, judge W. R. Overton in 1981 (McLean vs. Arkansas) and judge Jones in 2005 (Kitzmiller vs. Dover) based their decisions against the teaching of creationism in school on demarcation criteria between science and religious belief proposed by philosophers of science (cf. [Bird 1998](#), 2–9; [Branch 2011](#), 174). To tell the full truth, there were also philosophers of science, such as Laudan ([1983](#)), who have argued against the possibility of demarcating science from pseudoscience, and they have frequently been quoted by creationists as support for their position (to the embarrassment of other philosophers; cf. [Pennock 2011](#), 176ff).
- b. Equally significant is the critical role of philosophy of science concerning the commercial and political utilization of science, which counteracts the danger of *abusing* science and its results. Politicians, the media, and business leaders enjoy quoting expert opinions, which are presented all too often in a false or one-sided way in order to serve preordained purposes ([sec. 2.10.1](#), [4.3.6](#) – [7](#), [4.4](#)). But also applied scientists are often tempted to put subjective (e.g. financial) interests above objectivity. An example is the discovery of systematic biases in pharmaceutical research toward the financially supporting firms; philosophers of science have promoted scientific journal policies that may help to counteract such biases ([Brown 2008](#)).

1.2 On the Historical Development of Philosophy of Science

Although the term “philosophy of science” was only introduced in the 20th century, the philosophy of science is as old as science itself and had a parallel development (see [Losee 2001](#)). The history of the philosophy of science begins with *Aristotle* (384–322 B.C.), the ancient world's great systematizer of knowledge. Aristotle was much more oriented toward experience than was his teacher Plato. He was, nevertheless, like most philosophers after him, a follower of the *foundationalist* epistemological program. In this epistemological program, it is presumed that real knowledge is only possible if it stands on a foundation of certain and necessary principles, which are arrived at *not* through uncertain experience, but through rational intuition (cf. [Triplett 1990](#), sec. I; [Albert 1985](#), 18–28). Aristotle spoke in this context of “intuitive induction” ([Losee 2001](#), sec. Ia). In the current philosophy of science, by contrast, the *fallibilist* epistemological program has the upper hand, which concedes that our understanding of reality is basically *fallible*, and our scientific knowledge can be more or less well confirmed, but it cannot be guaranteed to be free of error. In a broader sense, of course, the scientific method is foundation-oriented, too, insofar it justifies its knowledge from most evident insights (primarily observation and logic) by inferences that are as reliable procedures possible – in contrast to coherentist epistemological programs, according to which the mutual coherence of beliefs is sufficient for their justification (cf. [Schurz 2024a](#), [sec. 2.1](#) – [2.2](#)).

Considerable advances in science and its philosophy had already been achieved by the late Middle Ages ([Crombie 1959](#)). Under the influence of philosophers such as *Roger Bacon* (1214–1292) or *William of Ockham* (1295–1349), the foundationalist program was gradually replaced by empirically oriented methodology. Of lasting importance is the principle of *Ockham's razor* , according to which one should reject all theoretical assumptions which are not needed to explain the observable. The decisive breakthrough, however, did not take place until the early modern period, the Philosophy of the Enlightenment. Building on the great successes of scientific method in this historical period – thanks to pioneers such as *Galileo Galilei* (1564–1642), *Isaac Newton* (1642–1727), and (later) *Charles Darwin* (1809–1882) – the influential *empiricist* movement was established, beginning with *Francis Bacon* (1561–1626) and *John Locke* (1632–1704) until *David Hume* (1711–1776) and *John Stuart Mill* (1806–1873). In the same period, the foundationalist epistemological program developed into the *rationalist* movement, prominently represented by *René Descartes* (1596–1650), *Gottfried W. Leibniz* (1646–1716), and *Immanuel Kant* (1724–1804).

The essential difference between the empiricist and rationalist philosophies can be characterized as follows. For the empiricists, sentences which may be justified *a priori* – that is, by pure reason and with rational certainty – are limited to so-called analytical sentences, that is, sentences whose truth is based on logic and conceptual convention. Such sentences possess no realistic content; they say nothing about the real world – only *synthetic* sentences may do this, but they can only be justified *a posteriori*, that is, based on experience. For the rationalists, on the other hand, *a priori* justifiable sentences with realistic content do exist, namely the so-called synthetic *a priori* sentences. However, both Descartes' and Kant's attempts to justify synthetic *a priori* principles of the sciences were unsuccessful and have been refuted by the further development of science. The development of the modern philosophy of science is closely bound with the insight that the foundationalist epistemological program was not tenable – at least not in its 'strong' version as explained above.

For a long time, classical systems of philosophical empiricism too had reductionist and foundationalist tendencies. The skeptical consequences of empiricism, which called for epistemic humility, were first recognized and radically formulated by Hume. Hume showed that the two central tenets of scientific method, the *causality* and the *induction principle* , are neither logically nor empirically justifiable. The causality principle says that all regular successions between events are produced by cause–effect relations; but Hume objected that all what one can observe are the successions of the events, while the idea that the earlier event or 'cause' produces the later event as its "effect" is metaphysical fiction (see [sec. 6.6.2](#)). The induction principle asserts that regular connections between events that have been observed in the past will also hold in the future; but Hume argued that this inference can neither be justified by logic nor by experience (see [sec. 2.11.1](#)). The problems involved in the justification of the principles of induction and causality have occupied philosophers up to the present day.

The apparent *insufficient justifiability* of empiricism put wind in the sails of subsequent rationalist movements, in which the foundationalist epistemological program, however, was gradually given up or replaced by pragmatized or historicized 'rationalisms' (see [sec. 1.3.2](#) –1.3.3). In the 20th century, *post-empiricist* and *post-rationalist* approaches have moved considerably closer to each other, and in the context of this tension, present-day philosophy of science has developed.

1.3 Philosophical Positions in Contemporary Philosophy of Science

1.3.1 Logical Empiricism

One of the most important origins of modern Philosophy of Science and Analytic Philosophy is logical empiricism, which was developed in particular by the Vienna Circle, but also by others, for example by members of the Berlin circle around Hans Reichenbach (1891–1953). The *Vienna Circle* was a group of philosophers and scientists in Vienna, whose core consisted of Moritz Schlick (1882–1936), Otto Neurath (1882–1945), and Rudolf Carnap (1891–1970). In the tradition of Ernst Mach (1838–1916), this group attempted to reestablish empiricism and scientific philosophy in general (cf. [Stadler 2001](#)). What was new about their situation was the development of *modern logic*, which had not been founded until the end of the 19th century. Modern logic made it possible to represent any language-based epistemological system with mathematical precision, and so the Vienna Circle hoped they had found at last the methodological equipment for a philosophy which could progress in a scientific manner (cf. [Schlick 1930/1931](#), 155f). In the middle of the 1930s, the Vienna Circle, with political leanings from liberal to socialist, had to disband under pressure from the National Socialists. The majority of its members emigrated, some via other countries, to the United States, where the logical-empirical movement, uniting with related Anglo-American directions, developed into Analytic Philosophy and gained a foothold in Europe in the postwar period.

Even today, many textbooks present logical empiricism as if it had represented a single dogmatic position; in reality, however, its proponents were highly capable of learning and continuously revised their positions. What today's philosophy of science can learn from logical empiricism is not so much any particular thesis but rather the high standards of conceptual and argumentative precision. In the phase up to 1935, the philosophy of science of logical empiricism was restricted by *positivism* and *reductionism*. In the later phase, the logical empiricists gradually rejected their restricted positions and replaced them with theses, to which the designations “empiricism” or “positivism” in the usual sense *no longer apply*. In one regard, logical empiricism has parted company with its classical predecessors early on: with the refutation of the classical empiricist thesis that at least ‘direct’ observation reports are infallible. In the so-called protocol sentence debate, Neurath and Carnap – but not Schlick – became convinced that even observation sentences such as “there is a table there” are in principle fallible. Empiricism, according to this new view, just meant that observation sentences have an epistemically favored role within the whole system of discovery (see [Carnap 1932/1933](#); [Neurath 1934](#), 113).

Empirical Reductionism, by contrast, was retained for quite some time. Classical empiricism advocated the following reductionist thesis: all ‘serious’ scientific concepts must be translatable via a chain of definitions into observational concepts. This program of reduction was adopted by logical empiricism in its early phase ([Carnap 1928a](#)). Later, however, the view prevailed that *theoretical concepts* such as “power,” “electrical field,” or “human character” *cannot* be defined by observational concepts, but rather go far beyond the immediately observable ([Hempel 1951](#); [Carnap 1956a](#)). In the so-called *standard* philosophy of science of the 1960s, the classical empiricist epistemological model had already broken down. What remained was a *minimal empiricism*, which consisted of the requirement that scientific theories must have empirical consequences, by means of which they can be tested.

A similarly modest empiricism is defended by contemporary empiricists such as van Fraassen (1980, [2002](#)). He does not only reject infallibilism and reductionism but also the *linguistic* orientation of logical empiricism. Instead, he defends a semantic conception of theories, in which theories are regarded not as sets of sentences, but as sets of conceptual models, similar as in structuralist philosophy of science (cf. [sec. 3.1.2](#)). Characteristic of van Fraassen's empiricism (1980, sec. 1) are two ingredients: first, a defense of the concept of theory-independent observation against Kuhn-followers (see [sec. 1.3.3](#)). And second, his rejection of scientific realism in favor of an instrumentalist position, according to which the theoretical (i.e. non-observable) part of a theory can only be reasonably said to be more or less *empirically adequate*, but not true or false. In this latter respect, van Fraassen is more restrictive than Carnap (e.g. 1928a) who always argued for ontological

pluralism , according to which both realism and instrumentalism are possible but epistemically undecidable practical attitudes toward scientific theories.

Two boundaries , which early logical empiricism had tried to maintain, became permeable in contemporary ‘minimal’ empiricism. *First* , the boundary between *science* and *metaphysics* became porous: it turned out that principles, which are of no empirical consequence *in isolation* , may give rise to new empirical consequences when joined with other theoretical sentences (see [sec. 5.4.2](#) ; [Hempel 1951](#) , 115; [Stegmüller 1969](#) ff, vol. II.1, 293–295). In this way, the *post-positivist* philosophy of science gained new access to the discussion of metaphysical principles such as realism and causality. Conversely, “inductive” branches of metaphysics emerged that sought to justify their principles not by pure thought but through empirical science (cf. [Hüttemann/Schurz 2025](#)). *Second* , the boundary between analytic and synthetic sentences became porous when it was recognized that the meaning of theoretical terms is not determined by isolated definitions, but by the respective background theory as a whole ([Carnap 1956a](#) , [1963](#)). In parallel fashion, Quine ([1951](#)) argued that in *natural* , empirically observed languages, no clear criterion exists by which one could undertake a demarcation between analytic and synthetic sentences.

As a consequence of this, analytic philosophy’s *linguistic* orientation (e.g. [Carnap 1950a](#)) was gradually given up. Linguistically oriented philosophers try to answer all philosophical questions by means of semantical concept analysis. But if a clear divide between (analytic) questions of meaning and (synthetic) questions of fact is missing, this exclusively linguistic orientation is no longer tenable. For this reason, Carnap ([1956b](#) , suppl. E) started to work on a system of pragmatics. Later on, Quine ([1974](#)) developed a *naturalistic* view, according to which science and philosophy form a continuum, the latter’s theses also having to be corroborated by experience.

1.3.2 Critical Rationalism

Also associated with the Vienna Circle was *Karl Popper* (1902–1994), the founder of so-called critical rationalism. From the beginning, Popper’s theory of science displayed the very non-reductionist tendencies that logical empiricism had accepted only after years in transition. According to Popper, scientific theories may go any distance beyond our experience, as long as they are empirically testable. Popper’s tests consist of *falsification attempts* , which should be as severe as possible. Popper supported his emphasis on falsification (instead of verification) by citing the *asymmetry* of *verification* and *falsification* of strictly universal hypotheses such as “all metals conduct electricity.” Such hypotheses cannot be verified by any finite number of observations, because they speak about a potentially infinite domain of instances – in our example, all metals in the past and in the future. However, these hypotheses can be falsified by a single observed exception ([Popper 1935](#) /2002, sec. I–IV).

For Popper, falsifiability was the decisive criterion of *demarcation* between science and speculation: scientific theories are not verifiable, but they must be falsifiable. Once falsified, they are *rejected* as false; but if they withstand manyfold falsification attempts, they are viewed as *corroborated* .

Popper’s falsificationism was also subjected to diverse criticism in the discussion which followed. Lakatos ([1970](#)) pointed out that scientific theories are almost never rejected on the basis of one single counterexample, but may be immunized against conflicts with recalcitrant data by means of modifications that are usually more or less *ad hoc* (see [sec. 5.6.1](#)). Popper’s *anti-inductionism* – that is, his thesis that science could progress without making any use of induction – also came in for harsh criticism (see [sec. 2.7.2](#)).

Popper’s critical rationalism is far removed from classical rationalism. Nowhere in Popper’s philosophy is it claimed that knowledge of reality can be grounded in non-empirical a priori intuition; on the contrary, Popper always rejected this. Thus, Popper’s philosophy is *post-rationalist* just as that of the late logical empiricists is *post-positivist* . Popper stressed more strongly than late logical

empiricism, however, that observation sentences are not only fallible, but also *theoretically biased* , so that the borderline between observational and theoretical terms cannot be sharply drawn (1935/2002, sec. V.30 addendum; new app. X). With this argumentation, critical rationalism opened its doors to the attacks from historical relativism, to which we turn now.

1.3.3 The Historical Account and Relativism

At the end of the 1950s, the historian of science Thomas Kuhn ([1962](#)) rattled the scientific world by presenting a fundamental attack on the standard philosophy of science. Factual science, said Kuhn, behaves differently from the normative ideals ‘prescribed’ by the philosophers of science. The alternative scientific model developed by Kuhn takes more of a historical-sociological than a logical-cognitive approach. According to Kuhn, scientific development takes place on the basis of so-called paradigms, such as the paradigm of classical (Newtonian) physics or (Darwin’s) theory of evolution. A Kuhnian paradigm contains at least the following three components: (i) very general theoretical principles or models, (ii) typical examples of successful applications, and (iii) methodological-normative assumptions (cf. [Kuhn 1977](#) ; [Hoyningen-Huene 1993](#) ; [Schurz 1998a](#) , [2014c](#)). The sociological correlate of a paradigm is a *scientific community* , a community of experts in the given field, who stick to a paradigm and work on its further development. According to Kuhn, a paradigm determines not only the fundamental principles, but also the major open problems and the *interpretation* of observed data – no, it even determines the observed data *themselves* , because, according to this view, *all observation is theory-laden* , and hence theoretically biased. For Kuhn, there is no observation which is independent of background theory or paradigm. Kuhn takes this strong thesis from Hanson ([1958](#)).

According to Kuhn, scientific development takes place in two successive phases, a *normal* and a *revolutionary* phase. Common acceptance of a paradigm makes the continuous growth of knowledge possible in the normal phase of a scientific discipline. When recalcitrant data, so-called anomalies, resist explanation by the given paradigm, the conflict is settled by more-or-less *ad hoc* modifications of current versions of theories within the accepted paradigm. Should such anomalies become more frequent, however, younger scholars will begin to look for a *new* paradigm. As soon as one is found, scientific development enters a revolutionary phase for a certain time, in which two paradigms struggle for supremacy. Kuhn’s examples in (1962) include the transition from Ptolemaic to Copernican astronomy and from Newtonian to Einsteinian physics. Kuhn claims that there are no common standards of rationality during a paradigm change and all empirical data are seen by the new paradigm in a new way. For this reason, the two competing paradigms are, according to Kuhn’s ‘notorious’ *incommensurability thesis* , rationally incomparable. So the struggle for supremacy does not take place in the form of a comparison of cognitive achievement, but rather in the form of a *power struggle* , in which supporters of the old paradigm eventually die out, as a result of which the new paradigm prevails, and a new normal scientific phase is heralded in.

The standard philosophy of science was unable to effectively counter the argument of the theory-dependence of scientific observations and its relativistic consequences. For all questions about the cognitive nature of observation had been banned by the narrowed perspective of the *logic* of science. For Popper, the question of what should be called an observation sentence was a question of pure convention (1935/2002, sec. V.30). The logical empiricists saw it in much the same way. It will be shown in [sec. 2.8.2](#) that a different picture results as soon as we take on an *empirical-cognitive* perspective. After all, if there were no difference between what experience *shows* , and what theory leads us to *expect* , then empirical science would be an exercise in *permanent self-deception* . Any theory of the empirical sciences which cannot explain this central difference must be extremely unsatisfying. Moreover, the loss of a viable criterion of empiricism would also undermine any credible distinction between real and fake facts. After right-wing movements began to criticize scientifically established facts, for example, concerning the climate crisis, as “leftist ideology,”

philosophers with left-leaning views tend to embrace once again the experience-based objectivity that was criticized by them some decades ago as an expression of “positivist ideology.”

Kuhn’s alternative model of science was the subject of much controversial discussion in the decades that followed. As well as an opposition to Kuhn, both a moderate and a more radical branch of Kuhn-followers was formed (cf. [Hoyningen-Huene 1993](#) , 212f). The more radical group gave relativism a strong impetus – most prominently in Feyerabend’s “Anarchist Epistemology” (1975), who with his maxim “anything goes” argued against methodological rules that are guiding or should guide science. The moderate Kuhnians, by contrast, aimed to separate Kuhn’s insights in the philosophy of science from the exaggerations of his “radical interpretation.” Kuhn himself later softened his incommensurability thesis and granted limited recognition to the possibility of rational comparisons between paradigms. ¹A moderate Kuhn-follower, and a Popper-follower at the same time, is Imre Lakatos and his conception of scientific research programs. Another moderate Kuhn-follower is Larry Laudan ([1981](#)), who argued that the empirical success of a scientific theory is not necessarily a reliable indicator of its truth. Such moderately Kuhnian developments have brought important insights to light for the philosophy of science (see [sec. 5.6](#) and 5.8). A further controversy centered around Kuhn’s thesis that in its normal scientific phases a scientific discipline is always dominated by a single paradigm. Scientists and philosophers of science have pointed out that, contrary to Kuhn’s thesis, disciplines outside of physics often exhibit a persistent coexistence of rival paradigms (cf. [Schurz/Weingartner 1998](#) ; [Kornmesser/Schurz 2014](#)).

1.3.4 Metaphysical and Scientific Realism

In reaction to the Kuhnian challenges of the logical-empirical model of science, many philosophers have suggested returning to the metaphysical approaches of realism and necessity, from which the philosophy of science in the early 20th century had tried to free itself. For example, it is a consequence of post-positivist philosophy of science that the term “mass” actually has different *meanings* in Newtonian and Einsteinian physics. This theory-dependence was also one of the main pillars of Kuhn’s incommensurability thesis. The early Putnam (e.g. 1975b, 219ff, 229–235, 240f), together with Kripke ([1972](#) , 55–59), developed the following argumentation against this consequence: scientific progress in the sense of a movement toward objective truth is only possible if the *reference* of scientific terms, that is, their relationship to reality, *is rigidly fixed* with metaphysical necessity, in the sense that this reference stays the same in all possible theories and across all possible worlds. Kripke and Putnam assumed the reference fixation to take place by acts of baptizing, together with a rigid sameness-relation between natural kinds, with the result that statements such as “water is H₂O” become true by “metaphysical necessity” (see also [Shoemaker 1980](#)). Fine ([1994](#)) also regards such identity relations as metaphysical necessities constituted by essential properties.

We agree with Kripke and Putnam that the meaning of natural kind terms such as “water” cannot be learned by an empirical definition (as in [Carnap’s 1956b](#) account), but by acts of ostension (“this is water”) grasping the right sameness-in-kind relation between water exemplars. But as a number of arguments have shown, scientific knowledge cannot determine sameness-in-kind relations and the reference of (theoretical) concepts with metaphysical necessity. ²Of course, classification systems for scientific (e.g. chemical or biological) kinds can be fixed *analytically* – that is, through definitional conventions (cf. [Sidelle 1989](#)). However, which classification system accurately reflects reality depends on the theoretical background knowledge ³which undergoes historical change ([Hendry 2008](#)). The idea that a necessary relationship between language and experience can be established independently of experience is in the end a regression toward a foundationalist epistemic model. In fact, the later Putnam ([1977](#) , [1982](#)) relinquished his earlier metaphysical realism by arguing as follows: if neither our empirical nor our theoretical knowledge can determine the reference of our concepts in a unique way, then *nothing* can do this – the consequence is an

unavoidable “indeterminacy of reference.” The same indeterminacy of reference has been worked out by Quine, but while Quine remained a naturalistic realist, the late Putnam adopted a more-or-less idealistic position that he called “internal realism.”

An account of necessity that is metaphysical in a weaker sense than the above ones is that of Armstrong ([1983](#)). According to him, laws of nature supervene on an assumed necessitation relation between “universals” (such as “water” and “H₂O”), which is, however, *contingent* – that is, these necessities may be different in different possible worlds ([sec. 6.5.6](#)). A similar middle ground between metaphysics and empiricism is Lewis ([1973](#) , [1986a](#)), who developed a possible worlds semantics for counterfactual and causal conditionals, but at the same time defends an elaborated Humean account to laws of nature, according to which laws supervene on the ‘mosaic’ of empirical facts ([sec. 6.6.5](#) ; cf. [Schrenk 2017](#) , [sec. 4.3](#) , [5.3](#)).

Like metaphysical realists, proponents of scientific realism also advocate for the convergence of science toward truth – but not on the basis of metaphysical postulates, but based on the empirical success of science. To support this position, scientific realists require robust arguments that allow them to infer the approximate truth of a theory from its empirical success. The most well-known argument of this kind is Putnam’s *no miracles* argument (1975a, 73). According to this argument, the empirical success of a theory would be a sheer miracle unless one assumes that the theory is at least approximately true. The no miracles argument has been subject to extensive criticism, notably by Laudan ([1981](#)), who, in his *pessimistic meta-induction* , argued that virtually all empirically successful theories of the past have eventually been shown to be false, whence it is inductively highly probable that our current best theories will suffer the same fate. Other philosophers of science, such as Boyd ([1984](#)) and Psillos ([1999](#)), have defended scientific realism against such anti-realist arguments, leading to a scientific realism debate that continues to this day (see [sec. 5.8.3](#) – [5.8.8](#)).

1.3.5 Naturalism and the Cognitive Turn

Modern naturalism goes back primarily to Quine ([1951](#) , [1960](#)), who rejected the analytic-synthetic distinction, thereby dissolving the clear boundary between scientific philosophy and empirical science. For Quine, naturalized epistemology becomes an empirical discipline, which is “scientifically investigating man’s acquisition of science” ([Quine 1974](#) , 3). The naturalistic program was emphatically taken up by a number of young epistemologists and philosophers of science (e.g. [Kitcher 1992](#) ; [Papineau 1993](#) ; [Kornblith 1994](#)).

Philosophical naturalism found expression in several philosophical programs. A first concretization of naturalism is *evolutionary epistemology* (e.g. [Campbell 1984](#)): many aspects of humans’ cognitive capacities and achievements can be explained by facts of evolution, in particular regarding perceptual knowledge. It has to be critically remarked, however, that naturalistic approaches themselves make certain epistemological assumptions which are *in need of justification* , for example concerning realism and induction, and so they cannot entirely replace classical epistemology. There is, however, an epistemological school of thought that indeed amounts to a *renunciation* of justification in this regard, namely epistemological externalism. According to its most well-known version ([Goldman 1986](#)), a belief counts as justified if it is produced by an “objectively reliable” method, without requiring anyone to be able to justify the reliability of the method in question. A justification of inductive prediction methods is not sought; instead, it is simply asserted that a predictive method is justified if it is reliable, that is, if its predictions will come out as true. Otherwise, it is considered unreliable. The externalist treatment of realism works in a similar way. However, since all nonscientific ‘methods’ – ranging from creationist doctrines of salvation to clairvoyance – also claim reliability for themselves, externalist epistemology fails to demarcate science from nonscience and thereby condemns itself to epistemic uselessness (see [Schurz 2024a](#) , [sec. 2.4](#)).

A second concretization of naturalism is the so-called cognitive turn. This turn resulted from the fact that meanwhile the cognitive sciences had attained the status of an interdisciplinary subject area in their own right, in which the basis of humans' cognitive and scientific capabilities is examined not only logically in philosophy but also empirically in psychology or linguistics and computationally in Artificial Intelligence research (cf. [Thagard 1996](#) ; [Anderson 2000](#)). As a result, several philosophers have started efforts to bring the approaches of the cognitive sciences into philosophy of science and vice versa (e.g. [Goldman 1986](#) ; [Bechtel 1988](#) ; [Bishop/Trout 2005](#)). The recent movement of “experimental philosophy” has even started to examine philosophical intuitions by empirical investigations instead of the traditional “armchair method,” with interesting results concerning cultural differences ([Machery et al. 2004](#) ; [Knobe/Shau 2008](#) ; [Horvath/Koch 2020](#)).

The third specification of naturalism is, finally, naturalistic philosophy of science. Like Quine, it assumes a continuity between the methods of the empirical sciences and those of the philosophy of science. Questions of value are not excluded from science, but they are understood in an instrumental way – as questions of the optimal choice of means for given ends ([Laudan 1987](#)) – which closely aligns with the value-neutrality approach discussed in [sec. 2.5](#) . By contrast, the influence of externalism must be viewed critically: the philosophical aim of justifying fundamental scientific methods in contrast to pseudoscience is largely abandoned, and the reliability of scientific methods and results is simply presupposed ([Godfrey-Smith 2003](#) , 151). This runs counter to the aim of theoretical clarification of science and the exposure of the ideological misuse of scientific authority, as outlined in [sec. 1.1](#) .

1.3.6 Structuralist and Semantic View of Theories

In search of a better defense against the Kuhnian offensive, Stegmüller ([1976](#) ; [1977](#) , vol. II., sec. 6; 1979b) suggested replacing the “statement view” of the standard philosophy of science with a so-called non-statement view. Stegmüller's suggestion is based on the idea, traceable to Suppes ([1957](#) , 246ff) and Sneed ([1971](#)), that it is better to reconstruct scientific theories as systems of set-theoretical *models* than as systems of statements in a logical language. Such systems of models are also called *structures* : hence, the name “structuralist” philosophy of science. In 1976, [Stegmüller \(1976](#) , 105) envisions this step as a radical upheaval. Later, he is more modest and argues for the transition from logical to set-theoretical formalization methods for “pragmatic and psychological reasons” concerning the simplicity of formal reconstructions (1979b, 65). For set-theoretical models are also characterized by statements, namely by statements in the language of set theory. Therefore, there exist straightforward possibilities of *translation* between the theory representations in the statement view and the non-statement view (see [sec. 3.1.2](#) ; [Schurz 2014b](#)).

While the statement-versus-non-statement debate turned out to be a bit of a red herring, the structuralist philosophy of science can claim significant achievements in the *detailed reconstruction* of individual scientific theories and theory nets (cf. [Balzer et al. 1987](#) ; [Balzer/Moulines 1996](#)). Moreover, the structuralist philosophy of science has received support from Anglo-Saxon philosophers such as van Fraassen (1980, sec. 2; 1989, sec. 9) and Giere ([1988](#) , [2004](#)), who called this position the “semantic view of theories” and developed it in a less formal way (cf. [Suppe 2000](#)). The semantic view of theories essentially consists in the claim that theories should be understood as models – a view later taken up by the “new mechanists” in the philosophy of science (e.g. [Glennan 2017](#) ; [Glennan et al. 2022](#) ; see [sec. 6.4.1](#)). However, the concept of a model remains vague: virtually anything can serve as a ‘model’ – from images to mathematical equations – as long as it serves to represent some aspect of reality (see [Glennan 2017](#) , 60; for criticism, see Chakravartty 2001; [Schurz 2014b](#)).

1.3.7 Pragmatic Philosophy of Science

In the 1980s, several authors spoke of a *pragmatic turn* in philosophy of science ([Stegmüller 1969](#) ff, vol. I, sec. XI; [Böhler et al. 1986](#) ; [Bernstein 2001](#)). Unfortunately, however, pragmatic philosophy and the use of the term “pragmatics,” in general, fall into two *opposing* camps. While one camp (e.g. Ch. S. [Peirce 1931](#) ff, van Fraassen 1980; [Rescher 1998](#) , [2001](#) ; [Schurz 1998](#) b) understands pragmatics in the *epistemically internal* sense, the other camp (e.g. [James 1907](#) ; [Rorty 1982](#) ; [Stich 1990](#) ; partly [Putnam 1995](#) ; Bernstein 2010) sees pragmatics in an *epistemically external* sense.

For the epistemically external camp, the pragmatic components of discovery are essential, but have nothing to do with its truth. By contrast, for the epistemically internal camp, it is these very pragmatic components, which are supposed to establish an objective conception of truth in a noncircular sense. The epistemically external camp makes knowledge (or justified belief) dependent on *any* kind of interests, be it predictive accuracy, political control, or emotional well-being, and sees therein a compelling argument for the unavoidable subjectivity of knowledge. In this sense, Stich ([1990](#) , 131ff) has argued that there are so many different conceptions of truth as there are different cognition-guiding interests. By contrast, for the internal camp, the reference to epistemically *internal* purposes is the very thing which makes possible a noncircular foundation of knowledge. For example, the reference to the cognitive purpose of maximizing true predictions enables a justification of induction that escapes Hume’s skeptical challenge ([sec. 2.11.2](#)). In conclusion, the external and the internal camp *differ so greatly* in their understanding of pragmatics, that they should not be subsumed under a single term.

1.3.8 Hermeneutics and Critical Theory

The *Philosophy of the Humanities* , which originated in the 19th century, goes back to Friedrich Schleiermacher (1768–1834), Wilhelm Dilthey (1833–1911), and Wilhelm Windelband (1848–1915). It started as a movement for the methodological demarcation of the humanities against the expanding natural sciences. Central to this philosophy is the use of *hermeneutics* , the study of interpersonal *understanding* (*Verstehen*) , for the purpose of such a methodological demarcation. The *methodological dualism* of the philosophers of the humanities has been brought to a head with the following thesis: in the natural sciences, we *explain* , but in the humanities, we *understand* . While scientific explanations are based on general law hypotheses, it was claimed that understanding in the humanities needs no recourse to such law hypotheses, because the human mind is not subject to strict law mechanisms, and understanding has direct access to the human mind.

Hermeneutics is part of a tradition from well before Schleiermacher, in which there was no mention of any rigid separation of the humanities from logic and natural sciences (see [Scholz 2001](#) , part I.B). The methodological dualism originated in the 19th century and had a strong influence on methodology debates of the 20th century (cf. [Von Wright 1971](#) ; [Manninen/Tuomela 1976](#) ; [Apel 1984](#)). The contemporary controversy in the humanities and social sciences about *quantitative* versus *qualitative* methods is partially a revivification of the old methodological dispute (cf. [Lamnek 1988](#) ; [Denzin/Lincoln 2000](#)). In response, the analytic philosophy of science has produced a number of approaches, in which the unity of hermeneutics and empirical methods is worked out (see [sec. 6.4.2](#) ; cf. [Hirsch 1973](#) ; [Follesdal 1979](#) ; [Taylor 1980](#) ; [Bühler 2003](#)).

Critical theory was developed by the philosophers of the so-called Frankfurt school, mainly Horkheimer, Adorno, Marcuse, and Habermas. It combines the methodologically dualist version of hermeneutics with the doctrine of the unavoidable interest-dependence of discovery and knowledge (see [Habermas 1966](#) ; [Apel 1984](#)). Critical theory also influenced the philosophy of science debate on the value-freedom versus value-ladenness of scientific research. We turn to this question in [sec. 2.5](#) and 2.10, where a modernized version of scientific value neutrality is developed – one that is compatible with value commitment in science and capable of withstanding relevant objections.

1.3.9 Radical Constructivism, Postmodernism, and the Science Wars

Radical constructivism was developed in the 20th century by biologists such as Maturana and Varela ([1992](#)) and psychologists such as von Glasersfeld ([1995](#)); philosophically, however, it is much older. The central argument of constructivism infers that, because our perceptions and descriptions of reality are constructed, there is no subject-independent reality – instead, *reality itself* is said to be a human construction. Even the early Kuhn embraced a constructivist position, claiming that “the world itself” has changed following a scientific paradigm shift ([Kuhn 1962](#) , chapter 10; see [Godfrey-Smith 2003](#) , 96).

In [sec. 2.8.2](#) , we will show that the argumentation of radical constructivism is based on a *fallacy* . Regardless of this, radical constructivism, in combination with the hermeneutic methodological dualism, has exerted its influence in the cultural and social sciences. There, it shaped French poststructuralism and the philosophy of postmodernism ([Lyotard 1984](#)), which reject scientific reality and its impact on human culture, instead promoting the idea that reality is entirely socially constructed. In the United States, these currents gave rise to the field of “Science Studies” and encouraged a left-leaning politicized understanding of science, as exemplified by feminist philosophy of science ([Harding 1986](#) ; cf. [Godfrey-Smith 2003](#) , sec. 9.1–9.4). Most pointedly, the social constructivist position can be found in the “science studies” of Latour and Woolgar ([1986](#)) and Knorr-Cetina ([1981](#)), where scientific laboratories (e.g. in nuclear physics or biochemistry) are studied by ethnographical methods of studying “foreign cultures” without engaging with the actual epistemic content being investigated, in order to learn something about the “social construction of scientific reality.”

The disregard of methodological standards and the misuse of scientific terminology met with strong opposition from the scientific community and led, in 1996, to a “counterattack” in the form of the *Sokal Affair* . In order to expose the lack of qualifications behind the wordplay of postmodern authors, Sokal ([1996](#)) published a hoax article titled “Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity,” which consisted of a jumble of quotes from postmodern thinkers. Despite its nonsensical content, the article was accepted and published by the postmodern-oriented American journal *Social Text* . When Sokal later revealed in another journal that his article had been a parody of postmodern discourse, the scandal was complete (cf. [Godfrey-Smith 2003](#) , 146f). The subsequently published book by Sokal and Bricmont ([1998](#)), entitled *Fashionable Nonsense: How the Postmodernists Abuse Science* , caused a major stir and became a catalyst for the so-called Science Wars ([Ross 1996](#)).

1.4 Introductory Literature and Exercises

1.4.1 Introductory Literature

A historically oriented introduction into the philosophy of science of the 20th and 21st centuries is Godfrey-Smith ([2003](#)). Ladyman’s introduction (2002) is focused on natural science and Rosenberg ([2008](#)) introduces to the philosophy of social science. Okasha ([2016](#)) gives a systematic but very brief introduction. An excellent introduction to the history of the philosophy of science is offered by Losee ([2001](#)). Stadler ([2001](#)) presents an extensive study of the Vienna Circle and its logical empiricism. For Popper’s philosophy of science, see [Keuth \(1998](#) , ed.). Hoyningen-Huene ([1993](#)) offers an introduction into the philosophy of Kuhn. [Hüttemann \(2021\)](#) offers an introduction into the metaphysics of science; for critical discussion of metaphysical accounts, see Ladyman and Ross ([2007](#)). On evolutionary epistemology, see Campbell ([1984](#)). The history and current situation of hermeneutics are dealt with by Bruns ([1992](#)).

1.4.2 Questions and Exercises

Sec. 1.1–1.2

1. What are the seven main questions in the general philosophy of science?
2. What is the major difference between the empiricist and rationalist ways of thinking in epistemology?

Sec. 1.3

1. What is understood by empirical reductionism?
2. What criterion of demarcation is suggested by critical rationalism?
3. What is a paradigm for Kuhn? Discuss Kuhn's incommensurability thesis.
4. How does scientific instrumentalism differ from scientific realism?
5. Discuss the methodological dualism of the philosophers of the humanities from a contemporary viewpoint.

Notes

1. Kuhn ([2000](#) , 33–57); cf. Okasha ([2016](#) , 87); [Seselja/Straßer \(2013](#) , 14f). Empirical data show that history of science proceeds much more cumulative than Kuhn suggests (see [Krauss 2025](#)). ↵
2. See [sec. 5.8.2](#) . Cf. Mellor ([1980](#) , 118f); Bird ([1998](#) , 108–120); Ladyman/Ross ([2007](#) , 290ff). ↵
3. In philosophy of science, “background knowledge” is not understood in the strict epistemological sense: it contains scientific beliefs that are rationally accepted but not necessarily true. ↵

2 The Question of Unity

In Search for Common Foundations of the Sciences

DOI: [10.4324/9781003675075-2](https://doi.org/10.4324/9781003675075-2)

This chapter has been made available under a CC BY-SA license.

2.1 Normative or Descriptive? The Method of Rational Reconstruction

Prima facie, we can identify two opposing views of the goals and methods of the philosophy of science:

1. According to the *normative* view, it is the goal of philosophy of science to say what science *should* be, and how it should be pursued. For this purpose, philosophy of science must determine in what scientific rationality consists, and on the basis of which criteria a scientific hypothesis may be justified.
2. According to the *descriptive* view, on the other hand, it is the task of philosophy of science to say what science de facto *is*, and how it proceeds. For this purpose, philosophy of science must describe and explain the existing sciences in terms of their historical development and their current structure.

The normative view is historically older, and currently still enjoys a wide following. Both logical empiricists and critical rationalists held this view. It was not until Kuhn ([1962](#)) set off the historical turn in philosophy of science that the descriptive counterpoint became popular. It can be found in the work of historically oriented philosophers of science (cf. [McMullin 1970](#) ; [Laudan 1977](#) , 162), of the structuralists (e.g. [Stegmüller 1976](#) , 267ff; [Balzer et al. 1987](#) , xvff), and of the naturalists (cf. [Giere 1999](#) , 160–163; [Ladyman 2002](#) , 4). So which position is right?

The position of the normativists can be traced back to Hans Reichenbach's ([1938](#) , 6f) distinction between the *context of discovery* , that is, the factual *genesis* of scientific knowledge, and its *context of justification* . The plausible prima facie argumentation of the normativists may be broken down as follows (e.g. [Popper 1935](#) /2002, [sec. 1.2](#)):

(2.1-1) *Normativist argumentation* :

Premise 1 : To assess the epistemic value of a hypothesis, *only* the question of rational justifiability of its truth claim is relevant.

Premise 2 : Questions about the hypothesis's historical development, the causes for its discovery, etc., are irrelevant to this purpose.

Conclusion : Therefore, philosophy of science should only examine science's rational justification and not its context of genesis – the latter being the remit of the history, sociology, or psychology of science, but having nothing to do with its philosophy.

Since Kuhn ([1962](#)), the argumentation of the normativists has been criticized in a very *sweeping* way. Here, we will attempt to clarify what is correct and what is flawed in this line of argument. The argument is certainly logically correct – assuming we take it to be analytically true that the philosophy of science is concerned with the assessment of the epistemic value of hypotheses. That is to say, *if* the premises of the argument were really true, then the conclusion would also be true. There is also *no error* in premise 1 – this thesis also seems to be analytically true, to the extent that science is concerned with the discovery of truth, and nothing else. But premise 2 is already infected by the 'germ' of the normativist fallacy. Because the

historical genesis of a scientific hypothesis includes not only so-called external factors which are epistemically irrelevant, but also internal factors which are epistemically relevant: for example, the experimental methods, which scientists de facto use, and the rational hypothesis evaluation, which they de facto apply.

The organic chemist Kekulé, for example, reported that his groundbreaking hypothesis of ring-shaped benzene molecules first appeared to him in a *dream* (see [Hempel 1966](#) , 16). This fact was, of course, epistemically irrelevant and played no role in the justification of the hypothesis. By contrast, the *Boyle-Mariotte law of gases* was discovered by means of the systematic extrapolation of data ([Langley et al. 1987](#) , 81) – and this fact played a decisive role in the justification of the hypothesis.

Philosophers should *learn* from the de facto methods and arguments of science as far as they are successful, and they can only do this by observing them. Instead of premise 2, all the normativists usually want to say, or at least all they should say, is the following *almost trivial* weakening:

(2.1–2) *Premise 2** : The epistemically *external* part of the history of the development of a hypothesis is irrelevant to the assessment of its epistemic value.

Of course, the conclusion of argument ([2.1-1](#)) no longer follows if premise 2 is replaced by the weaker premise 2*. And with this conclusion, the error of the normativist argumentation becomes *grave* . To realize this, we must simply ask the following question: how can philosophers of science *acquire* the appropriate rules and criteria for the scientific method? It would be naïve to presume that these rules and criteria might be gained through pure logic or mere intuition; science is much too complex for that. That would be like somebody trying to write a bridge-building manual without having examined the de facto practice of bridge building and typical examples of existing bridges. The philosophy of science, then, *can* and *should* learn from the real examples provided by the history of science. This brings us to the following result: although it is the *primary* task of the philosophy of science to explain the methods and criteria of *rational* science, it can only fulfill this task by *also* studying the *factual sciences* in their historical development.

Kuhn's criticism crystallized into the insight that the scientific models provided by logical empiricism and critical rationalism were too *simplicistic* to give a faithful picture of real science (cf. [Stegmüller 1979b](#) , 5; 1969ff, vol. II.3, 20). Reacting to this, young theorists proclaimed that the philosophy of science should even be restricted to the descriptive analysis of the sciences. But this position too is overstated. The question of the definition and criteria of scientific rationality must, of course, remain central to philosophy of science, as it is this question which defines philosophy of science as a discipline and sets it apart from the history and sociology of science. In summary, the philosophy of science is a discipline, which contains both descriptive and normative elements – an insight that has been worked out by [Stegmüller \(1969](#) ff, vol. IV.1, 9ff).

The philosophy of science's method may best be described as a *rational reconstruction* . My understanding of this method is different from Lakatos' understanding of it, which has been sharply criticized by Laudan ([1977](#) , 167–170). For Lakatos ([1971](#) , 119f), 'history' as understood by rational reconstruction is not the *actual* history, but a projection of ideals of rationality into history, and a selection of those parts of history that conform to that ideal. By contrast, in my understanding, the actual history forms one pole of the method.

More precisely, rational reconstruction moves between two poles: a (so-called) *descriptive corrective* , which should be adequately reconstructed, and a *normative corrective* , which contains rational norms. The descriptive corrective does not contain every piece of factual sciences. It contains only typical examples of de facto successful (i.e. unanimously and sustainably accepted) scientific discoveries. Moreover, it also contains typical 'counterexamples,' that is, examples of de facto refuted scientific hypotheses. Philosophy of science bases its reconstructions on these consensually accepted examples of scientific success or failure, though only in a preliminary and defeasible way.

The normative corrective consists of certain epistemic intuitions that are normally assumed by scientists, in spite of their various cognitive differences. That such a common ground exists is the position that will be defended in this chapter. First of all, the normative corrective contains one supreme epistemic goal or norm, which may be explicated as follows (cf. [Weingartner 1978](#) , [sec. 3.2](#)):

(Prop. 2.1-1) The *supreme epistemic goal* (G) of science is to find *true* and *content-rich* statements, laws, or theories, relating to a given domain of phenomena.

A sentence is the richer in content, the more *consequences* it has (see [sec. 3.7](#)). The restriction of the goal of science to true *and* content-rich statements is significant, because the probability and the content of hypotheses are often *inversely proportional*. We can maximize the truth chances of hypotheses by sticking to trivial tautologies like: “The Sun revolves around the Earth, unless it doesn’t.” Inversely, it is easy to present very substantial and impressive hypotheses if we disregard their truth value, for example, “I have invented a perpetual motion machine.” The true art of the scientist is to formulate hypotheses which have both empirical truth and substantial content and consequences.

In the above explication of the supreme goal of science, there appears the little word “truth.” One might object that the goal of truth is not metaphysically neutral. However, we understand this concept in a liberal way, leaving room for different metaphysical interpretations (see [sec. 2.8.1](#)). On the other hand, we don’t want to be too liberal about the meaning of “truth.” For people have rather different ideas of what truth is. While one person seeks truth in the platonic heaven, another person may equate truth with empirical success, and for a third one truth is what makes herself happy. To ensure that the goal of scientific discovery as explicated in [prop. 2.1-1](#) is not entirely vague concerning the meaning of “truth,” we need a *minimal* explication of what “search for content-rich truths” is intended to mean. This is provided by the second ‘normative’ component that is normally assumed by scientists as well as by philosophers of science, namely a minimal common epistemological model of scientific discovery. In [sec. 2.2](#), we will explain this minimal epistemological model of the sciences in detail. The key points are given here: it contains (1) a *minimal realism*, (2) *fallibilism* and critical attitude, (3) the struggle for *objectivity*, (4) a *minimal empiricism*, and (5) the struggle for *logical clarity*. Although this minimal epistemological model consists primarily not of normative, but of *descriptive* claims, we still regard it as part of the normative corrective of philosophy of science, as it gives clear meaning to the supreme goal of science, and is usually seen as a precondition of science.

Given the above described normative and descriptive correctives, the method of rational reconstruction may be precisely characterized as follows:

(Def. 2.1-1) *Method of rational reconstruction* : Develop generalized and logically precise *models* of scientific discovery, which on the one hand *fit the descriptive corrective*, that is, provide optimal explanations for the success of the examples and the failure of the counterexamples from the history of science, and which on the other hand may be *justified according to the normative corrective*, as providing optimal *means* of reaching the supreme goal of scientific discovery, given the assumptions of the minimal epistemological model.

In short, the philosophy of science abstracts generalized models from the descriptive corrective, which at the same time are (or should be) in line with the normative corrective. This is summarized in the schematic diagram in [fig. 2.1-1](#).

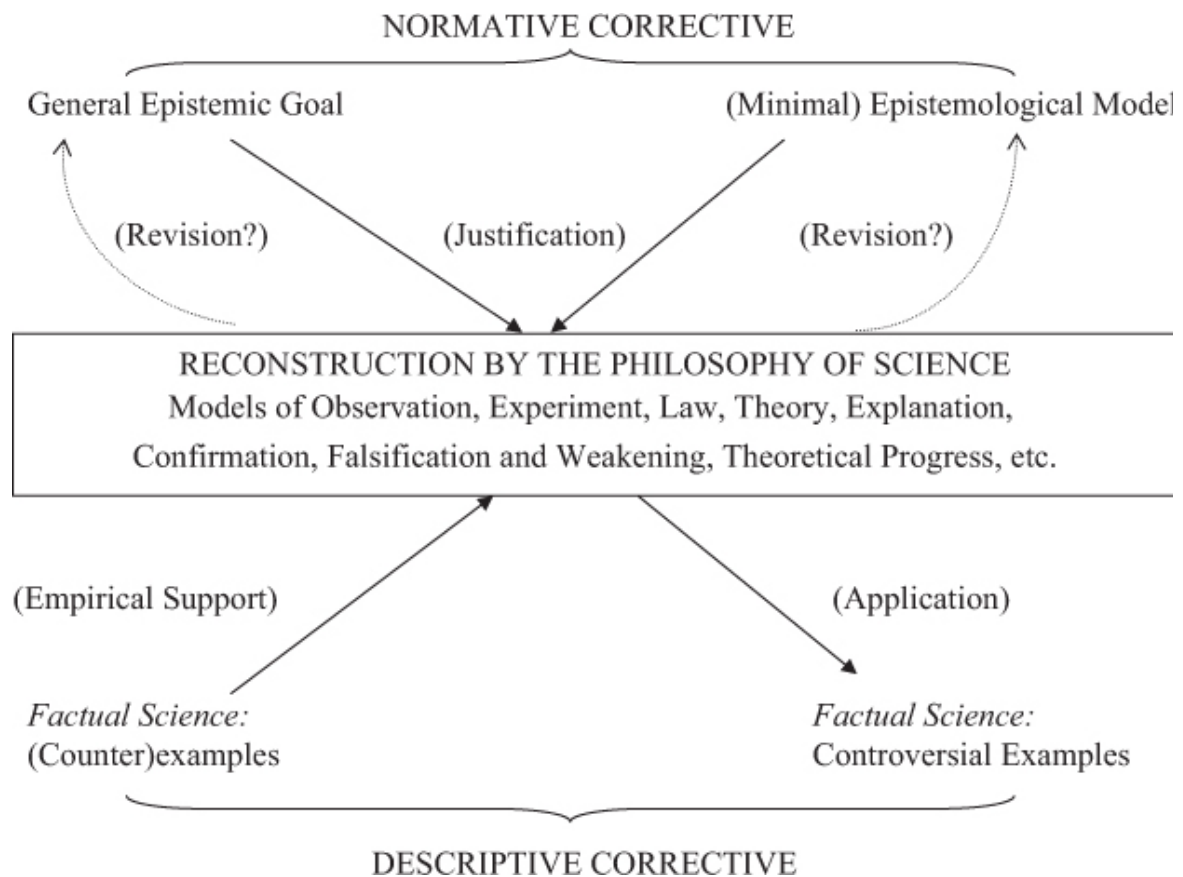


Figure 2.1-1 The method of rational reconstruction. [↗](#)

As long as both aspects of rational reconstruction – descriptive adequacy and normative justifiability – can be made to work in harmony with each other, the work of scientific reconstruction can be counted as successful and its results as *corroborated*. Of course, we cannot expect that everything we find in *factual science* will also be *rational* according to the normative corrective. For this reason, the descriptive corrective of the philosophy of science contains only *accepted* (counter) examples. When models developed by the philosophy of science have been corroborated, we can take the next step of applying them to *controversial* areas of scientific disciplines, in which no consensus has been reached – be it the objectivity problem of quantum mechanics, the teleology problem in biology, or the explanation-understanding problem in the humanities. It is exactly here that philosophy of science can supply valuable decision-making aids to the individual sciences. So, on the one hand, the philosophy of science learns from the individual sciences and is corrected by them; on the other hand, its reconstruction work can be instructive to the individual sciences and thus help them solve specific problems.

The goal of rational reconstruction, though, bringing together descriptive analysis and normative justification, does not always have to succeed. This applies, for example, when philosophers of science proceed from unrealistically high or one-sided normative standards, which will persistently disallow them to reconcile their normative corrective with the descriptive analysis of science. In such cases, the philosophy of science must go on to *revise* its own normative corrective. When it does this – this has already happened several times in the history of the philosophy of science – then the philosophy of science, to borrow Kuhn's terminology, is no longer in a normal scientific phase, but in a revolutionary phase.

To conclude this section, we will compare the model of rational reconstruction with two similar characterizations of the method of philosophy of science. Carnap ([1950b](#) , sec. 2) described the procedure of the philosophy of science as *explication*. Here, a vague concept of everyday or scientific language (the explicandum), for example, the idea of probability, is replaced by a precisely defined term (the explicatum), whereby the explicatum must differ from the explicandum insofar as it eliminates vagueness and ambiguity,

but it should also remain as close to it as it can and be as fruitful and simple as possible. Brun ([2016](#)) describes Carnap's method of explication as "conceptual re-engineering." Carnap does not explain, however, how exactly these partially contradictory demands should be combined. In comparison, the proposed model of rational reconstruction contains clear normative and descriptive guidelines; and by distinguishing normal and revolutionary phases in the philosophy of science, it takes into account the possibility that incoherencies can arise between the two. There are also some similarities between rational reconstruction and the method of reflective equilibrium according to Goodman ([1955](#) , 85–89) and Rawls ([1971](#) , 20f, 48–52). However, while the method of reflective equilibrium works purely *coherently* , aiming at the reciprocal adaptation of methodological rules and intuition, rational reconstruction is *empirically grounded* by its descriptive corrective. Moreover, the normative corrective consists not of mere intuition, but of epistemological *hypotheses* , together with a supreme epistemological norm.

2.2 Common Epistemological Assumptions

The minimal epistemological model, common to more or less all scientific disciplines, contains five epistemological assumptions (E1–E5):

E1 – Minimal realism : According to this assumption, there is a reality, which exists independently of the (given) epistemological subject. It is not assumed that all properties of this reality are knowable. The possibility of fundamental limitations to scientific knowledge is left open and cannot be answered a priori, but only based on the factual epistemic success of science. Scientific disciplines aim to provide statements which are as true and as content-rich as possible about clearly defined areas of this reality. The term "truth" is used here in the sense of the *structural correspondence theory* , according to which the truth of a sentence consists in its structural agreement with the part of reality described by it. This notion of structural correspondence, which was first precisely defined by Alfred Tarski ([1936a](#)), does *not* imply any intrinsic relationship or mirroring between language and reality. It also allows that theories which are strictly speaking *false* may be true in an *approximative* sense, or *close* to the truth ([sec. 5.6.2](#) , 5.11.1).

E2 – Fallibilism and critical attitude : There is no infallible 'golden path' to the truth in the sense of the correspondence theory. According to the assumption of fallibilism, every scientific statement is more or less fallible; so, we can never be *absolutely* sure of their truth. Fallibility goes hand in hand with a critical attitude, according to which no statement may ever be once and for all exempted from criticism. However, we can regard the truth of a statement as more or less probable, and in the limiting case as "practically certain," in which case we consider its error probability negligible. For this reason, everything depends on using empirical tests to find out more about the *probability* of a scientific hypothesis, in the sense of assumption E4 below.

E3 – Objectivity and intersubjectivity : According to this assumption, a statement's truth must hold objectively, that is, independently of the beliefs and value-attitudes of the epistemological subject, because by assumption E1, reality has subject-independent existence, and truth is the correspondence between statement and reality. So, the assumption of objectivity already follows from assumption E1. However, the characterization of objectivity as subject-independence does not help us in the practice of science, as only subjects make statements and formulate hypotheses. We have no direct access to objective truth; we can only approach it indirectly via – not certain, but probable – *criteria* .

The second part of assumption E3 says that a central criterion for objectivity and also for truth lies in the *intersubjectivity* of scientific statements. Intersubjectivity, correctly understood, is a plausible consequence as well as a plausible indicator of objectivity and truth. It is a plausible consequence because if a statement's truth can be established at all convincingly, then – at least in principle – it must be possible to convince any person of the statement's truth, provided the person is sufficiently competent, that is, cognitively trained and informed. And it is a plausible indicator, because the agreement of many mutually *independent* and competent (i.e. fallible but truth-biased) researchers on the hypothesis provably increases its truth chances. This is an instance of the *Condorcet jury theorem* (cf. [Bovens/Hartmann 2003](#) , [sec. 5.2](#) ; [Dietrich/Spiekermann 2023](#)).

E4 – Minimal empiricism : With the exception of formal sciences such as mathematics, which we will discuss later, the objects examined by any science must in principle be accessible to experience and observation. After all, reliable information about reality can only be gained through perceptual observation; for perception is our only form of direct informational contact with reality. Empirical observations, then, are a decisive referee in the scientific search for truth: scientific law hypotheses and theories must be *tested* in terms of them. This is the central statement of assumption E4. Observations are not infallible (that would contradict assumption E2), but they offer the easiest and fastest route to intersubjectivity and practical certainty. Observations are “epistemically privileged,” as one also says. The empiricism contained in assumption E4 is *minimal*, because different from empirical reductionism it is not claimed that all scientific statements must be traceable back to observations via a chain of definitions, or must be provable by observations. Scientific theories can and should contain theoretical concepts, by which they can talk about things which lie beyond what is observable to human senses. The decisive requirement of E4 is merely that statements about the unobservable must have *observable consequences* by which they may be tested.

E5 – Logic in the wider sense : Using precise logical methods to introduce terms, to formulate sentences, and to construct correct arguments, is the most effective way to approach the goal of content-rich truth (according to assumptions E1–E4). For the meaning of a sentence cannot be established until the terms used in it have been clearly defined. And only sentences with a precisely formulated meaning may be examined for their logical consequences. Finally, only after the consequences of a hypothesis are known, may this hypothesis be precisely empirically tested in accordance with assumption E4. In conclusion, the procedure of empirical testing requires the use of exact logical methods at all times – in the wider sense of logic, not restricted to deductive logic.

Our minimal epistemological model makes as few assumptions as possible, but as many as necessary, in order to explain the possibility of objective science. Some of its assumptions are worthy of further explanation. We call the realism in assumption E1 “minimal” because it only asserts the objective *existence* of reality. Whether and to what extent reality is objectively discoverable cannot, according to assumption E2, be answered a priori, but only through the success of scientific discovery methods. This sets our minimal realism apart from metaphysical realism, in all its varieties. Moreover, we don’t claim that science is impossible without the assumption of a subject-independent reality, but only that this assumption is *normally* made in scientific disciplines (see [sec. 2.8.1](#)).

To avoid misunderstandings, we should add that the correspondence theory of truth, while well suited as a *definition* of truth, does not offer any *criteria* to evaluate the truth of a sentence (cf. [Rescher 1973](#) , ch. 1; [Kirkham 1992](#) , 195). The path to such criteria is illuminated by assumptions E5 and E4: investigate the empirical consequences of your hypothesis and test them using empirical observations. For example, to find out the truth of the sentence: “this flower is red,” the correspondence-theoretical insight that this sentence is true if the flower is red will not get me anywhere; rather, I must be in a position to apprehend the object denoted by the ostensive individual term “this flower” and the perceptual quality expressed by the predicate “red” visually, in order to test the truth of the sentence by a simple act of perception.

In assumptions E4 and E5, we claim the existence of only two minimal levels, in which discovery finds its origin and its touchstone: observation, which allows us fragmentary but reliable and intersubjective access to reality, and logic in its wider sense, which provides a universal cognitive tool, but in itself says nothing about reality.

In assumption E3, we limited our characterization of *intersubjectivity* to “sufficiently competent and informed” persons. This limitation is necessary, because persons who do not *understand* the content of a hypothesis, or who do not possess the relevant *evidence* , are unable to test a hypothesis for its truth. This difficulty shows us that intersubjectivity can on no account be used as a *definition* of objectivity or truth. The latter suggestion has been made in the constructivist school of Kamlah and Lorenzen ([1984](#) , 119), and one century earlier by the founder of the philosophical school of *pragmati(ci)sm* , Charles S. Peirce ([1878b](#)). But competence is an uncertain and gradual criterion, meaning that any number of persons judged to be competent may still err collectively. Peirce attempted to save his consensus definition by equating truth with the consensus that an *ideal* community of researchers with an infinite amount of time would approach in the limit. But even in this counterfactual version, an intersubjective consensus does not have to agree with the

truth of correspondence theory. There are countless questions on whose true answer even an ideal community of researchers can never achieve consensus, because there is no evidential access to them, although there is no reason to assume that these propositions have an indefinite truth value – for example, “How many microbes were located in Caesar’s nose as he crossed the Rubicon?” In conclusion, intersubjectivity can at most be an uncertain criterion for objectivity.

2.3 Common Methodological Features

From the supreme epistemic goal (G) and the five epistemological assumptions (E1–5), there follow very plausibly four methodological features (M1–4), which are common to all empirical sciences in the broad sense. First, a word on terminology. By an *observation sentence*, we understand a sentence that expresses an *observable* state of affairs. We distinguish between *actual* and *potential* observation sentences. We call an observation sentence actual if the state of affairs expressed by it has indeed *been observed*. By contrast, an observation sentence is merely potential, if the state of affairs expressed by it has not been observed, but could be observed, for example, at a later time or in a different place. Merely potential observation sentences express empirical *predictions*, for example, “there will be a solar eclipse tomorrow.” We can now formulate the four common methodological features as follows (similarly in [Weingartner 1978](#), [sec. 3.1](#) – [3.7](#)):

- M1* : Science searches for hypotheses which are as general and as content-rich as possible and recorded in a scientific language. In all disciplines, these hypotheses include *laws* and *theories*, and in some (e.g. the historical sciences), they also include *hypothetical singular sentences*.
- M2* : Science searches for *actual observation sentences*, as many as possible (and as relevant as possible), which reflect the results of observations, experiments, or measurements.
- M3* : Science attempts, with the help of general and hypothetical sentences, to *explain* the currently known actual observation sentences, and to *predict* new, and as yet unknown, *potential* observation sentences.
- M4* : Science attempts to *test empirically* its general and hypothetical sentences by comparing the predicted (potential) observation sentences with the currently known (actual) observation sentences. If the latter are in agreement with the former, the prediction was successful (it then becomes a successful explanation) and the hypothetical sentence (law or theory) is *confirmed* or *corroborated*. If there is a contradiction between the latter and the former, the prediction was without success and the law or theory is *falsified*, or in the case of a merely statistical prediction, *weakened*.

We only speak of an *explanation* if the state of affairs to be explained has already come to pass and is known, while in a *prediction*, the predicted state of affairs is new and unknown. Viewed logically, both explanations and predictions take the form of deductive or probabilistic arguments (see [sec. 6.1](#) – [6.4](#)).

These four methodological features or modules are closely connected with the epistemological assumptions of science. Feature M1 results from the epistemic goal G, because laws and theories are the typical examples of general, content-rich sentences. M2 results from assumption E4 (sharpened by assumptions E2 and E3): to discover more about the conjectured truth of its hypotheses, science must collect empirical data. In order for general hypotheses to be tested on the basis of these data, the logical or probabilistic consequences of the hypotheses, in the form of predictions or potential explanations, must be investigated, and these consequences compared with the data – in this way, the methodological building blocks M3 and M4 are motivated by the assumptions E4 and E5.

Successful predictions and explanations not only have a bearing on the ultimate scientific goal of finding true and content-rich theories, but they are also of crucial importance for human *planning behavior*, indeed, to our survival. Predictions allow people to plan their future by *adapting themselves* to the predicted sequence of events. Explanations are concerned with finding causes ([sec. 6.2](#)); they provide people with *knowledge of causes*. Causal knowledge gives people more power than merely predictive knowledge: it enables them to causally intervene in the course of events, and thereby *adapt* the world to suit *their own wishes*. The scientific subgoals of prediction and causal control have been pursued by humanity since the Stone Age. Anyone who can *predict* that nourishing plants are growing in a given area can decide to migrate there and thereby ensure a high reproduction rate for their own tribal group. But those who know the *causes* of the growth of nourishing plants, who know how to plant and cultivate them, need no longer be gatherers, but may become farmers, and practice agriculture instead.

In advanced science, the scientific procedures of explanation and prediction operate not on two but on three levels: actual observation sentences at the lowest level, empirical law hypotheses at the middle level, and scientific theories at the upper level. Between the first and the second, and between the second and the third level, the methodological procedures from M3 and M4 take place, as shown in [fig. 2.3-1](#) . This threefold division of scientific work is advantageous for many reasons; in all advanced scientific disciplines, a division of labor has arisen between practical experimenters (levels 1 and 2) and theoreticians (levels 2 and 3). To give an example: In the case of the physical theory of the planetary orbits ([sec. 5.1](#)), level 1 corresponds to the actual astronomic observations (e.g. by Tycho Brahe), level 2 corresponds to the empirical laws of the (elliptical) planetary orbits around the Sun (inductively obtained by Johannes Kepler), and level 3 to Isaac Newton's theoretical explanations of these orbits by the Sun's gravitational attraction together with the planets' centrifugal force.

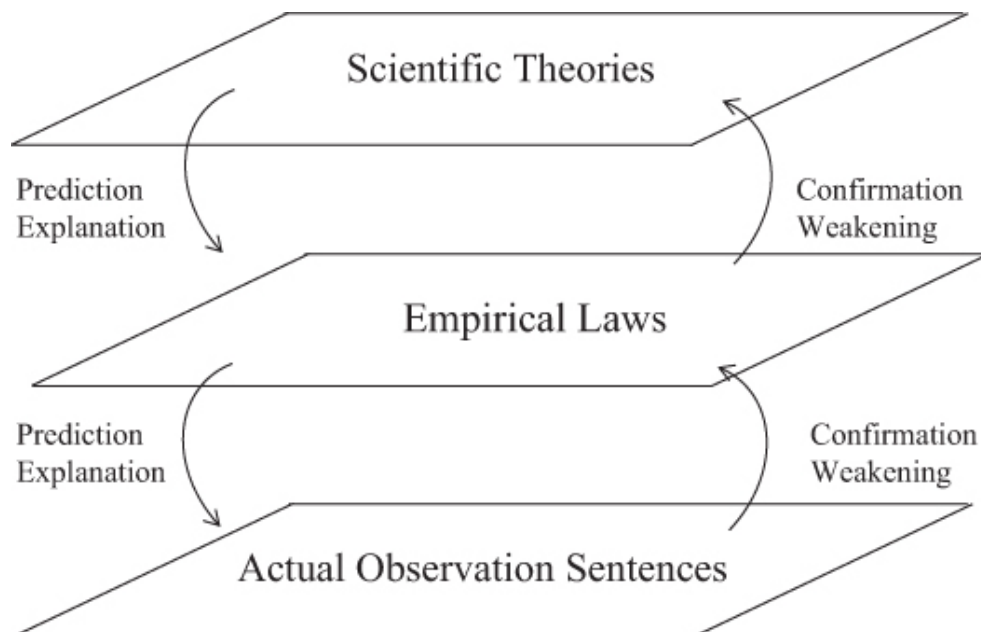


Figure 2.3-1: The three levels of the scientific method. [↗](#)

2.4 Scientific Disciplines and Their Classification

In this chapter, we will examine to what extent our epistemological assumptions E1–5 and the methodological features M1–4 match the existing scientific disciplines. To structure this investigation systematically, we introduce something that is rarely found in textbooks on the philosophy of science: a division of scientific disciplines according to their area of inquiry (see 2.4-1). Our division is certainly imperfect and we attribute to it not a final but merely a heuristic status. Some disciplines are mixed disciplines and are listed under “also;” further special disciplines can be added:

(2.4-1): *Scientific disciplines, classified according to area of inquiry :*

Sciences...

1. ... *of nature* : physics, chemistry, biology, geology, medicine (astronomy, cosmology, geography, paleontology, history of biological evolution)
2. ... *of technology* : mechanical and electrical engineering ..., also: computer science
3. ... *of human beings* : psychology (also: education, medicine, cognitive science)
4. ... *of (human) society* : sociology, economics, political science (also: anthropology, ethnology, geography)
5. ... *of (human) history* : history (also: anthropology, ethnology, philosophy as history of ideas)

6. ... *of human cultural (mental, social) artifacts* : legal sciences, linguistics, literary science, sciences of fine arts and music, art history, media studies (also: education, religious studies)
7. ... *of formal structures (formal sciences)* : mathematics (logic, statistics, theoretical computer science, system theory...), formal methodology and philosophy of science
8. ... *of the general foundations of human ideas* : philosophy (epistemology, philosophy of science and theoretical philosophy; ethics, esthetics, and practical philosophy)
9. ... *of God* : theology (also: religious studies)

Our classification has the advantage of not being tainted by partly value-based categories such as (natural) *science*, *humanities*, or *real science* (etc.). If we try to apply these catch-all categories to our area-of-inquiry-based classification, we recognize that there is no clear consensus about what goes where, and that different systems of classification often have more to do with the politics than the philosophy of science. Group 1, for example, covers the core *natural sciences*. But mathematics too sees itself as a natural science, although it belongs to the category of *formal sciences*; moreover, in the meantime, mathematics is not only applied in natural sciences but also in economics and social sciences. Of late, psychology also sees itself predominantly as a natural science, although it could equally be assigned to the humanities or social sciences.

Groups 5 and 6 comprise the core of the traditional *humanities*. Philosophy, though, is just as close to the natural sciences as it is to the humanities, although in contemporary universities, it is usually placed in the faculty of arts and humanities. In [sec. 1.3.8 – 1.3.9](#), we explained that since the 19th century, the traditional humanities were increasingly understood as a counterpart or even an opposition to the natural sciences. The term *humanities and social sciences* was coined in the 1960s and intended to serve as a modernized counterpart to the natural sciences. Whether this terminological group covers only groups 3 and 4, or also groups 5 and 6, was not terminologically fixed.

In the last two decades, the new terminological group of the *cultural sciences* was formed, in part as a reaction to contemporary tendencies in psychology and social sciences to increasingly use methods of natural sciences. The term “cultural sciences” is predominantly used in continental Europe and is understood in a broader sense than the Anglo-Saxon “cultural studies,” though its extension is not precisely defined. The frequently supposed opposition between natural and cultural sciences is an instance of what Snow ([1960](#)) has called the two ‘cultures’ in the current scientific age. In this book, we intend to refute the view that there exists a principal methodological *gap* between the sciences of nature and culture; this gap, where it exists, is more a matter of different value-attitudes and worldviews than a matter of unbridgeable differences in method.

In the term *factual sciences* (in German “Realwissenschaft”), all sciences are brought together that have a part of the *real* world as an area of inquiry; so, factual sciences include all groups apart from the formal sciences. Theology can only be seen as a factual science if we assume the reality of God, as this reality cannot be scientifically demonstrated, its inclusion here is questionable.

2.4.1 The Special Status of Mathematical Sciences

First, we have to explain the special status of group 7, mathematics, and formal sciences. In their totality, assumptions E1–5 and features M1–4 only make sense for the factual-empirical sciences. Mathematics and formal sciences do not have a domain of inquiry that consists of *concrete* and empirically accessible objects, but rather deal with formal and *abstract* structures. Mathematics, then, has no empirical basis, and therefore assumption E5 and methodological features M2 and M4, which require an empirical basis, do not apply. This does not mean that mathematics is a speculative discipline, which makes statements about reality with no empirical basis: pure mathematics makes no statements at all about an (empirically accessible) *reality*. All it does is develop formal structures and models, characterize them by using abstract axioms and definitions which are as general as possible, and derive logical consequences from them. It would be nonsense, for example, to say that the axiomatic system of real numbers or of Euclidean geometry was per se “true” or “false.” Such claims can only be made when mathematical terms are interpreted empirically, for example, when real locations or time points are assigned to real numbers, or real distances to mathematical vectors. But then geometry is no longer a mathematical theory, but becomes a physical one. Einstein once formulated it like this: “As far as the theorems of mathematics relate to reality, they are not certain; and as far as they are certain, they do not relate to reality” (cf. [Carnap 1967](#), 56). So, we can say: exactly *because* mathematics

ignores the empirical interpretation of its models, it is the strictest of all sciences. To be sure, mathematics concerns itself primarily with those structural models which are applicable to certain areas of reality. Seen in this way, mathematical disciplines can be more or less *fruitful* for the real sciences, but they cannot be true or false in the sense of the factual sciences.

Assumptions E2–4 and the methodological feature M1 can be applied to mathematics without restriction (the assumption of fallibility E2 insofar as even mathematical proofs are not guaranteed to be error-free). Assumption E1 of minimal realism applies only in an indirect way, insofar as mathematical structures have the purpose of being applied to the factual sciences. The methodological component M3 applies insofar as mathematicians are interested in the derivation of the various theorems, which express important properties of their models. Such derivations may be termed mathematical *explanations*, as long as it is clear that these are not causal explanations of real facts (cf. [Weingartner 1978](#), 102; [Pincock 2015](#); [Kuorikoski 2021](#)).

Let us not neglect the fact that the *philosophy of mathematics* includes opposing views on the nature of mathematical objects (cf. [Benacerraf/Putnam 1964](#); [Shapiro 2000](#); [Linnebo 2020](#)). According to the widespread view which is shared by us, mathematical objects are abstract *conceptual* objects, that is, conceptual constructions. Real objects, for example, five red roses, can *exemplify* mathematical objects, for example, the number five, but can never be identical with them. This view, called *conceptual structuralism*, was worked out in detail by the group of mathematicians around Bourbaki ([1961](#)) and philosophers of mathematics like Field ([1980](#)). The opposing standpoint in the philosophy of mathematics is *mathematical realism*, which assumes that mathematical objects exist *per se*. This position has been defended, for example, by Frege ([1918](#), 43) and Gödel ([1947](#), 271f). Also in this position, mathematical entities are seen as abstract and *platonic* objects, not located in the reality of space and time, but in a “platonic world of mathematical ideas” which is accessible not empirically, but only through mathematical “intuition.” An argument in favor of mathematical structuralism is Ockham’s razor (cf. 5.8.7): structuralism explains mathematical phenomena equally well as realism, but without the assumption of strange platonic entities. Another position is mathematical naturalism ([Maddy 1997](#)), according to which mathematical theories resemble scientific theories because they can fit reality more or less well in their applications. (The analogous view in the philosophy of logic has been called “anti-exceptionalism;” cf. [Hjortland 2017](#).) Against this view, it must be objected that in mathematical theories the reference to factual applications is fundamentally left open, whereas in empirical theories this reference is semantically fixed. As a result, empirical theories can be tested through experiments, which is undoubtedly *not* the case for mathematical theories.

In summary, mathematics and the other formal sciences are a *specialization* within the whole system of science, dealing with formal models, algorithms, and methods. This specialization has shown itself to be particularly *useful* in the history of science. Put in a nutshell, the information that the mathematician passes to the empirical scientist is this: “*if* your empirical domain satisfies the axioms of this-and-this mathematical structure, *then* it must also satisfy these-and-these theorems about this structure that we have proved; we grant you the satisfaction of the theorems, provided *you* grant us the satisfaction of the axioms.”

So, we characterize the formal sciences as *prerequisite and auxiliary sciences*, which furnish the factual sciences with formal foundations and structural models with known properties.

2.4.2 Testing the Common Foundations of Science

Having clarified the special status of mathematics, let us now ask: with the *exception* of the formal sciences, to what degree do the assumptions E1–5 and the methodological features M1–4 comply with the listed groups of disciplines? Obviously, E1–5 and M1–4 fit all disciplines in groups 1–3 (natural sciences, technical disciplines, psychology). There is also no reason not to apply them to the social sciences and historical disciplines in groups 4 and 5. Minimal realism, fallibilism, objectivity and intersubjectivity (E1–3), and logical clarity (E5) are accepted standards in all these disciplines, and the same is true for the search for general hypotheses, explanations, and predictions (M1, M3). In all these disciplines, empirical data are available, for example, in the form of the results of psychological experiments, interview data, economic and demographic data, and in the historical sciences in the form of historical texts and source materials, etc. By this, assumption E4 and building blocks M2 and M4 are also fulfilled. Of course, it is significantly more difficult to find general law hypotheses in social science and history, because of the complexity of their area of inquiry – but all that is required for the applicability of the methodological building blocks to a given