

## Scientific Objectivity

*First published Mon Aug 25, 2014; substantive revision Fri Oct 30, 2020*

Scientific objectivity is a property of various aspects of science. It expresses the idea that scientific claims, methods, results—and scientists themselves—are not, or should not be, influenced by particular perspectives, value judgments, community bias or personal interests, to name a few relevant factors. Objectivity is often considered to be an ideal for scientific inquiry, a good reason for valuing scientific knowledge, and the basis of the authority of science in society.

Many central debates in the philosophy of science have, in one way or another, to do with objectivity: confirmation and the problem of induction; theory choice and scientific change; realism; scientific explanation; experimentation; measurement and quantification; statistical evidence; reproducibility; evidence-based science; feminism and values in science. Understanding the role of objectivity in science is therefore integral to a full appreciation of these debates. As this article testifies, the reverse is true too: it is impossible to fully appreciate the notion of scientific objectivity without touching upon many of these debates.

The ideal of objectivity has been criticized repeatedly in philosophy of science, questioning both its desirability and its attainability. This article focuses on the question of how scientific objectivity should be *defined*, whether the ideal of objectivity is *desirable*, and to what extent scientists can *achieve* it.

- [1. Introduction](#)
- [2. Objectivity as Faithfulness to Facts](#)
  - [2.1 The View From Nowhere](#)
  - [2.2 Theory-Ladenness and Incommensurability](#)
  - [2.3 Underdetermination, Values, and the Experimenters' Regress](#)
- [3. Objectivity as Absence of Normative Commitments and the Value-Free Ideal](#)
  - [3.1 Epistemic and Contextual Values](#)
  - [3.2 Acceptance of Scientific Hypotheses and Value Neutrality](#)
  - [3.3 Science, Policy and the Value-Free Ideal](#)
- [4. Objectivity as Freedom from Personal Biases](#)
  - [4.1 Measurement and Quantification](#)
  - [4.2 Statistical Evidence](#)
    - [4.2.1 Bayesian Inference](#)
    - [4.2.2 Frequentist Inference](#)
  - [4.3 Feyerabend: The Tyranny of the Rational Method](#)

- [5. Objectivity as a Feature of Scientific Communities and Their Practices](#)
    - [5.1 Reproducibility and the Meta-Analytic Perspective](#)
    - [5.2 Feminist and Standpoint Epistemology](#)
  - [6. Issues in the Special Sciences](#)
    - [6.1 Max Weber and Objectivity in the Social Sciences](#)
    - [6.2 Contemporary Rational Choice Theory](#)
    - [6.3 Evidence-based Medicine and Social Policy](#)
  - [7. The Unity and Disunity of Scientific Objectivity](#)
  - [8. Conclusions](#)
  - [Bibliography](#)
  - [Academic Tools](#)
  - [Other Internet Resources](#)
  - [Related Entries](#)
- 

## 1. Introduction

Objectivity is a value. To call a thing objective implies that it has a certain importance to us and that we approve of it. Objectivity comes in degrees. Claims, methods, results, and scientists can be more or less objective, and, other things being equal, the more objective, the better. Using the term “objective” to describe something often carries a special rhetorical force with it. The admiration of science among the general public and the authority science enjoys in public life stems to a large extent from the view that science is objective or at least more objective than other modes of inquiry. Understanding scientific objectivity is therefore central to understanding the nature of science and the role it plays in society.

If what is so great about science is its objectivity, then objectivity should be worth defending. The close examinations of scientific practice that philosophers of science have undertaken in the past fifty years have shown, however, that several conceptions of the ideal of objectivity are either questionable or unattainable. The prospects for a science providing a non-perspectival “view from nowhere” or for proceeding in a way uninformed by human goals and values are fairly slim, for example.

This article discusses several proposals to characterize the idea and ideal of objectivity in such a way that it is both strong enough to be valuable, and weak enough to be attainable and workable in practice. We begin with a natural conception of objectivity: **faithfulness to facts**. We motivate the intuitive appeal of this conception, discuss its relation to scientific method and discuss arguments challenging both its attainability as well as its desirability. We then move on to a second conception of objectivity as **absence of normative commitments and value-freedom**, and once more we contrast arguments in favor of such a conception with the challenges it faces. A third conception of objectivity which we discuss at length is the idea of **absence of personal bias**.

Finally there is the idea that objectivity is anchored in **scientific communities and their practices**. After discussing three **case studies** from economics, social science and medicine, we address the **conceptual unity of scientific objectivity**: Do the various conceptions have a common valid core, such as promoting trust in science or minimizing relevant epistemic risks? Or are they rivaling and only loosely related accounts? Finally we present some conjectures about what aspects of objectivity remain defensible and desirable in the light of the difficulties we have encountered.

## 2. Objectivity as Faithfulness to Facts

The basic idea of this first conception of objectivity is that scientific claims are objective in so far as they faithfully describe facts about the world. The philosophical rationale underlying this conception of objectivity is the view that there are facts “out there” in the world and that it is the task of scientists to discover, analyze, and systematize these facts. “Objective” then becomes a success word: if a claim is objective, it correctly describes some aspect of the world.

In this view, science is objective to the degree that it succeeds at discovering and generalizing facts, abstracting from the perspective of the individual scientist. Although few philosophers have fully endorsed such a conception of scientific objectivity, the idea figures recurrently in the work of prominent twentieth-century philosophers of science such as Carnap, Hempel, Popper, and Reichenbach.

## 2.1 The View From Nowhere

Humans experience the world from a perspective. The contents of an individual's experiences vary greatly with his perspective, which is affected by his personal situation, and the details of his perceptual apparatus, language and culture. While the experiences vary, there seems to be something that remains constant. The appearance of a tree will change as one approaches it but—according to common sense and most philosophers—the tree itself doesn't. A room may feel hot or cold for different persons, but its temperature is independent of their experiences. The object in front of me does not disappear just because the lights are turned off.

These examples motivate a distinction between qualities that vary with one's perspective, and qualities that remain constant through changes of perspective. The latter are the objective qualities. Thomas Nagel explains that we arrive at the idea of objective qualities in three steps (Nagel 1986: 14). The first step is to realize (or postulate) that our perceptions are caused by the actions of things around us, through their effects on our bodies. The second step is to realize (or postulate) that since the same qualities that cause perceptions in us also have effects on other things and can exist without causing any perceptions at all, their true nature must be detachable from their perspectival appearance and need not resemble it. The final step is to form a conception of that "true nature" independently of any perspective. Nagel calls that conception the "view from nowhere", Bernard Williams the "absolute conception" (Williams 1985 [2011]). It represents the world as it is, unmediated by human minds and other "distortions".

This absolute conception lies at the basis of scientific realism (for a detailed discussion, see the entry on [scientific realism](#)) and it is attractive in so far as it provides a basis for arbitrating between conflicting viewpoints (e.g., two different observations). Moreover, the absolute conception provides a simple and unified account of the world. Theories of trees will be very hard to come by if they use predicates such as "height as seen by an observer" and a hodgepodge if their predicates track the habits of ordinary language users rather than the properties of the world. To the extent, then, that science aims to provide explanations for natural phenomena, casting them in terms of the absolute conception would help to realize this aim. A scientific account cast in the language of the absolute conception may not only be able to explain why a tree is as tall as it is but also why we see it in one way when viewed from one standpoint and in a different way when viewed from another. As Williams (1985 [2011: 139]) puts it,

[the absolute conception] nonvacuously explain[s] how it itself, and the various perspectival views of the world, are possible.

A third reason to find the view from nowhere attractive is that if the world came in structures as characterized by it and we did have access to it, we could use our knowledge of it to ground predictions (which, to the extent that our theories do track the absolute structures, will be borne out). A fourth and related reason is that attempts to manipulate and control phenomena can similarly be grounded in our knowledge of these structures. To attain any of the four purposes—settling disagreements, explaining the world, predicting phenomena, and manipulation and control—the absolute conception is at best sufficient but not necessary. We can, for instance, settle disagreements by imposing the rule that the person with higher social rank or greater experience is always right. We can explain the world and our image of it by means of theories that do not represent absolute structures and properties, and there is no need to get things (absolutely) right in order to predict successfully. Nevertheless, there is something appealing in the idea that factual disagreements can be settled by the very facts themselves, that explanations and predictions grounded in what's really there rather than in a distorted image of it.

No matter how desirable, our ability to use scientific claims to represent facts about the world depends on whether these claims can unambiguously be established on the basis of evidence, and of evidence alone. Alas, the relation between evidence and scientific hypothesis is not straightforward. [Subsection 2.2](#) and [subsection 2.3](#)

will look at two challenges of the idea that even the best scientific method will yield claims that describe an aperspectival view from nowhere. [Section 5.2](#) will deal with socially motivated criticisms of the view from nowhere.

## 2.2 Theory-Ladenness and Incommensurability

According to a popular picture, all scientific theories are false and imperfect. Yet, as we add true and eliminate false beliefs, our best scientific theories become more *truthlike* (e.g., Popper 1963, 1972). If this picture is correct, then scientific knowledge grows by gradually approaching the truth and it will become more objective over time, that is, more faithful to facts. However, scientific theories often change, and sometimes several theories compete for the place of the best scientific account of the world.

It is inherent in the above picture of scientific objectivity that observations can, at least in principle, decide between competing theories. If they did not, the conception of objectivity as faithfulness would be pointless to have as we would not be in a position to verify it. This position has been adopted by Karl R. Popper, Rudolf Carnap and other leading figures in (broadly) empiricist philosophy of science. Many philosophers have argued that the relation between observation and theory is way more complex and that influences can actually run both ways (e.g., Duhem 1906 [1954]; Wittgenstein 1953 [2001]). The most lasting criticism, however, was delivered by Thomas S. Kuhn (1962 [1970]) in his book “The Structure of Scientific Revolutions”.

Kuhn’s analysis is built on the assumption that scientists always view research problems through the lens of a paradigm, defined by set of relevant problems, axioms, methodological presuppositions, techniques, and so forth. Kuhn provided several historical examples in favor of this claim. Scientific progress—and the practice of normal, everyday science—happens within a paradigm that guides the individual scientists’ puzzle-solving work and that sets the community standards.

Can observations undermine such a paradigm, and speak for a different one? Here, Kuhn famously stresses that **observations are “theory-laden”** (cf. also Hanson 1958): they depend on a body of theoretical assumptions through which they are perceived and conceptualized. This hypothesis has two important aspects.

First, the *meaning* of observational concepts is influenced by theoretical assumptions and presuppositions. For example, the concepts “mass” and “length” have different meanings in Newtonian and relativistic mechanics; so does the concept “temperature” in thermodynamics and statistical mechanics (cf. Feyerabend 1962). In other words, Kuhn denies that there is a theory-independent observation language. The “faithfulness to reality” of an observation report is always mediated by a theoretical *überbau*, disabling the role of observation reports as an impartial, merely fact-dependent arbiter between different theories.

Second, not only the observational concepts, but also the *perception* of a scientist depends on the paradigm she is working in.

Practicing in different worlds, the two groups of scientists [who work in different paradigms, J.R./J.S.] see different things when they look from the same point in the same direction. (Kuhn 1962 [1970: 150])

That is, our own sense data are shaped and structured by a theoretical framework, and may be fundamentally distinct from the sense data of scientists working in another one. Where a Ptolemaic astronomer like Tycho Brahe sees a sun setting behind the horizon, a Copernican astronomer like Johannes Kepler sees the horizon moving up to a stationary sun. If this picture is correct, then it is hard to assess which theory or paradigm is more faithful to the facts, that is, more objective.

The thesis of the theory-ladenness of observation has also been extended to the **incommensurability of different paradigms or scientific theories**, problematized independently by Thomas S. Kuhn (1962 [1970]) and Paul Feyerabend (1962). Literally, this concept means “having no measure in common”, and it figures prominently in arguments against a linear and standpoint-independent picture of scientific progress. For instance, the Special Theory of Relativity appears to be more faithful to the facts and therefore more objective

than Newtonian mechanics because it reduces, for low speeds, to the latter, and it accounts for some additional facts that are not predicted correctly by Newtonian mechanics. This picture is undermined, however, by two central aspects of incommensurability. First, not only do the observational concepts in both theories differ, but the principles for specifying their meaning may be inconsistent with each other (Feyerabend 1975: 269–270). Second, scientific research methods and standards of evaluation change with the theories or paradigms. Not all puzzles that could be tackled in the old paradigm will be solved by the new one—this is the phenomenon of “Kuhn loss”.

A meaningful use of objectivity presupposes, according to Feyerabend, to perceive and to describe the world from a specific perspective, e.g., when we try to verify the referential claims of a scientific theory. Only *within* a peculiar scientific worldview, the concept of objectivity may be applied meaningfully. That is, scientific method cannot free itself from the particular scientific theory to which it is applied; the door to standpoint-independence is locked. As Feyerabend puts it:

our epistemic activities may have a decisive influence even upon the most solid piece of cosmological furniture—they make gods disappear and replace them by heaps of atoms in empty space. (1978: 70)

Kuhn and Feyerabend’s theses about theory-ladenness of observation, and their implications for the objectivity of scientific inquiry have been much debated afterwards, and have often been misunderstood in a social constructivist sense. Therefore Kuhn later returned to the topic of scientific objectivity, of which he gives his own characterization in terms of the shared cognitive values of a scientific community. We discuss Kuhn’s later view in [section 3.1](#). For a more thorough coverage, see the entries on [theory and observation in science](#), [the incommensurability of scientific theories](#) and [Thomas S. Kuhn](#).

## 2.3 Underdetermination, Values, and the Experimenters’ Regress

Scientific theories are tested by comparing their implications with the results of observations and experiments. Unfortunately, neither positive results (when the theory’s predictions are borne out in the data) nor negative results (when they are not) allow unambiguous inferences about the theory. A positive result can obtain even though the theory is false, due to some alternative that makes the same predictions. Finding suspect Jones’ fingerprints on the murder weapon is consistent with his innocence because he might have used it as a kitchen knife. A negative result might be due not to the falsehood of the theory under test but due to the failing of one or more auxiliary assumptions needed to derive a prediction from the theory. Testing, let us say, the implications of Newton’s laws for movements in our planetary system against observations requires assumptions about the number of planets, the sun’s and the planets’ masses, the extent to which the earth’s atmosphere refracts light beams, how telescopes affect the results and so on. Any of these may be false, explaining an inconsistency. The *locus classicus* for these observations is Pierre Duhem’s *The Aim and Structure of Physical Theory* (Duhem 1906 [1954]). Duhem concluded that there was no “crucial experiment”, an experiment that conclusively decides between two alternative theories, in physics (1906 [1954: 188ff.]), and that physicists had to employ their expert judgment or what Duhem called “good sense” to determine what an experimental result means for the truth or falsehood of a theory (1906 [1954: 216ff.]).

In other words, there is a gap between the evidence and the theory supported by it. It is important to note that the alleged gap is more profound than the gap between the premisses of *any* inductive argument and its conclusion, say, the gap between “All hitherto observed ravens have been black” and “All ravens are black”. The latter gap could be bridged by an agreed upon rule of inductive reasoning. Alas, all attempts to find an analogous rule for theory choice have failed (e.g., Norton 2003). Various philosophers, historians, and sociologists of science have responded that theory appraisal is “a complex form of value judgment” (McMullin 1982: 701; see also Kuhn 1977; Hesse 1980; Bloor 1982).

In [section 3.1](#) below we will discuss the nature of the value judgments in more detail. For now the important lesson is that if these philosophers, historians, and sociologists are correct, the “faithfulness to facts” ideal is untenable. As the scientific image of the world is a joint product of the facts and scientists’ value judgments, that

image cannot be said to be aperspectival. Science does not eschew the human perspective. There are of course ways to escape this conclusion. If, as John Norton (2003; ms.—see Other Internet Resources) has argued, it is material facts that power and justify inductive inferences, and not value judgments, we can avoid the negative conclusion regarding the view from nowhere. Unsurprisingly, Norton is also critical of the idea that evidence generally underdetermines theory (Norton 2008). However, there are good reasons to mistrust Norton's optimism regarding the ineliminability of values and other non-factual elements in inductive inferences (Reiss 2020).

There is another, closely related concern. Most of the earlier critics of "objective" verification or falsification focused on the relation between evidence and scientific theories. There is a sense in which the claim that this relation is problematic is not so surprising. Scientific theories contain highly abstract claims that describe states of affairs far removed from the immediacy of sense experience. This is for a good reason: sense experience is necessarily perspectival, so to the extent to which scientific theories are to track the absolute conception, they must describe a world different from that of sense experience. But surely, one might think, the evidence itself is objective. So even if we do have reasons to doubt that abstract theories faithfully represent the world, we should stand on firmer grounds when it comes to the evidence against which we test abstract theories.

Theories are seldom tested against brute observations, however. Simple generalizations such as "all swans are white" are directly learned from observations (say, of the color of swans) but they do not represent the view from nowhere (for one thing, the view from nowhere doesn't have colors). Genuine scientific theories are tested against experimental facts or phenomena, which are themselves unobservable to the unaided senses. Experimental facts or phenomena are instead established using intricate procedures of measurement and experimentation.

We therefore need to ask whether the results of scientific measurements and experiments can be aperspectival. In an important debate in the 1980s and 1990s some commentators answered that question with a resounding "no", which was then rebutted by others. The debate concerns the so-called "experimenter's regress" (Collins 1985). Collins, a prominent sociologist of science, claims that in order to know whether an experimental result is correct, one first needs to know whether the apparatus producing the result is reliable. But one doesn't know whether the apparatus is reliable unless one knows that it produces correct results in the first place and so on and so on *ad infinitum*. Collins' main case concerns attempts to detect gravitational waves, which were very controversially discussed among physicists in the 1970s.

Collins argues that the circle is eventually broken not by the "facts" themselves but rather by factors having to do with the scientist's career, the social and cognitive interests of his community, and the expected fruitfulness for future work. It is important to note that in Collins's view these factors do not necessarily make scientific results arbitrary. But what he does argue is that the experimental results do not represent the world according to the absolute conception. Rather, they are produced jointly by the world, scientific apparatuses, and the psychological and sociological factors mentioned above. The facts and phenomena of science are therefore necessarily perspectival.

In a series of contributions, Allan Franklin, a physicist-turned-philosopher of science, has tried to show that while there are indeed no algorithmic procedures for establishing experimental facts, disagreements can nevertheless be settled by *reasoned judgment* on the basis of *bona fide* epistemological criteria such as experimental checks and calibration, elimination of possible sources of error, using apparatuses based on well-corroborated theory and so on (Franklin 1994, 1997). Collins responds that "reasonableness" is a social category that is not drawn from physics (Collins 1994).

The main issue for us in this debate is whether there are any reasons to believe that experimental results provide an aperspectival view on the world. According to Collins, experimental results are co-determined by the facts as well as social and psychological factors. According to Franklin, whatever else influences experimental results other than facts is not arbitrary but instead based on reasoned judgment. What he has not shown is that reasoned judgment guarantees that experimental results reflect the facts alone and are therefore aperspectival in any interesting sense. Another important challenge for the aperspectival account comes from feminist epistemology

and other accounts that stress the importance of the construction of scientific knowledge through epistemic communities. These accounts are reviewed in [section 5](#).

### 3. Objectivity as Absence of Normative Commitments and the Value-Free Ideal

In the previous section we have presented arguments against the view of objectivity as faithfulness to facts and an impersonal “view from nowhere”. An alternative view is that science is objective to the extent that it is *value-free*. Why would we identify objectivity with value-freedom or regard the latter as a prerequisite for the former? Part of the answer is empiricism. If science is in the business of producing empirical knowledge, and if differences about value judgments cannot be settled by empirical means, values should have no place in science. In the following we will try to make this intuition more precise.

#### 3.1 Epistemic and Contextual Values

Before addressing what we will call the “value-free ideal”, it will be helpful to distinguish four stages at which values may affect science. They are: (i) the choice of a scientific research problem; (ii) the gathering of evidence in relation to the problem; (iii) the acceptance of a scientific hypothesis or theory as an adequate answer to the problem on the basis of the evidence; (iv) the proliferation and application of scientific research results (Weber 1917 [1949]).

Most philosophers of science would agree that the role of values in science is contentious only with respect to dimensions (ii) and (iii): the **gathering of evidence** and the **acceptance of scientific theories**. It is almost universally accepted that the choice of a research problem is often influenced by interests of individual scientists, funding parties, and society as a whole. This influence may make science more shallow and slow down its long-run progress, but it has benefits, too: scientists will focus on providing solutions to those intellectual problems that are considered urgent by society and they may actually improve people’s lives. Similarly, the proliferation and application of scientific research results is evidently affected by the personal values of journal editors and end users, and little can be done about this. The real debate is about whether or not the “core” of scientific reasoning—the gathering of evidence and the assessment and acceptance scientific theories—is, and should be, value-free.

We have introduced the problem of the underdetermination of theory by evidence above. The problem does not stop, however, at values being required for filling the gap between theory and evidence. A further complication is that these values can conflict with each other. Consider the classical problem of fitting a mathematical function to a data set. The researcher often has the choice between using a complex function, which makes the relationship between the variables less *simple* but fits the data more *accurately*, or postulating a *simpler* relationship that is less *accurate*. Simplicity and accuracy are both important cognitive values, and trading them off requires a careful value judgment. However, philosophers of science tend to regard value-ladenness in this sense as benign. **Cognitive values** (sometimes also called “epistemic” or “constitutive” values) such as predictive accuracy, scope, unification, explanatory power, simplicity and coherence with other accepted theories are taken to be indicative of the truth of a theory and therefore provide reasons for preferring one theory over another (McMullin 1982, 2009; Laudan 1984; Steel 2010). Kuhn (1977) even claims that cognitive values define the shared commitments of science, that is, the standards of theory assessment that characterize the scientific approach as a whole. Note that not every philosopher entertains the same list of cognitive values: subjective differences in ranking and applying cognitive values do not vanish, a point Kuhn made emphatically.

In most views, the objectivity and authority of science is not threatened by cognitive values, but only by **non-cognitive** or **contextual values**. Contextual values are moral, personal, social, political and cultural values such as pleasure, justice and equality, conservation of the natural environment and diversity. The most notorious cases of improper uses of such values involve travesties of scientific reasoning, where the intrusion of contextual values led to an intolerant and oppressive scientific agenda with devastating epistemic and social consequences. In the Third Reich, a large part of contemporary physics, such as the theory of relativity, was condemned

because its inventors were Jewish; in the Soviet Union, biologist Nikolai Vavilov was sentenced to death (and died in prison) because his theories of genetic inheritance did not match Marxist-Leninist ideology. Both states tried to foster a science that was motivated by political convictions (“Deutsche Physik” in Nazi Germany, Lysenko’s Lamarckian theory of inheritance and denial of genetics), leading to disastrous epistemic and institutional effects.

Less spectacular, but arguably more frequent are cases where research is biased toward the interests of the sponsors, such as tobacco companies, food manufacturers and large pharmaceutical firms (e.g., Resnik 2007; Reiss 2010). This **preference bias**, defined by Wilholt (2009) as the infringement of conventional standards of the research community with the aim of arriving at a particular result, is clearly epistemically harmful. Especially for sensitive high-stakes issues such as the admission of medical drugs or the consequences of anthropogenic global warming, it seems desirable that research scientists assess theories without being influenced by such considerations. This is the core idea of the

**Value-Free Ideal (VFI):** Scientists should strive to minimize the influence of contextual values on scientific reasoning, e.g., in gathering evidence and assessing/accepting scientific theories.

According to the VFI, scientific objectivity is characterized by absence of contextual values and by exclusive commitment to cognitive values in stages (ii) and (iii) of the scientific process. See Dorato (2004: 53–54), Rupy (2006: 190) or Biddle (2013: 125) for alternative formulations.

For value-freedom to be a reasonable ideal, it must not be a goal beyond reach and be attainable at least to some degree. This claim is expressed by the

**Value-Neutrality Thesis (VNT):** Scientists can—at least in principle—gather evidence and assess/accept theories without making contextual value judgments.

Unlike the VFI, the VNT is not normative: its subject is whether the judgments that scientists make are, or could possibly be, free of contextual values. Similarly, Hugh Lacey (1999) distinguishes three principal components or aspects of value-free science: impartiality, neutrality and autonomy. **Impartiality** means that theories are solely accepted or appraised in virtue of their contribution to the cognitive values of science, such as truth, accuracy or explanatory power. This excludes the influence of contextual values, as stated above. **Neutrality** means that scientific theories make no value statements about the world: they are concerned with what there is, not with what there should be. Finally, scientific **autonomy** means that the scientific agenda is shaped by the desire to increase scientific knowledge, and that contextual values have no place in scientific method.

These three interpretations of value-free science can be combined with each other, or used individually. All of them, however, are subject to criticisms that we examine below. Denying the VNT, or the attainability of Lacey’s three criteria for value-free science, poses a challenge for scientific objectivity: one can either conclude that the ideal of objectivity should be rejected, or develop a conception of objectivity that differs from the VFI.

### 3.2 Acceptance of Scientific Hypotheses and Value Neutrality

Lacey’s characterization of value-free science and the VNT were once mainstream positions in philosophy of science. Their widespread acceptance was closely connected to Reichenbach’s famous distinction between **context of discovery** and **context of justification**. Reichenbach first made this distinction with respect to the epistemology of mathematics:

the objective relation from the given entities to the solution, and the subjective way of finding it, are clearly separated for problems of a deductive character [...] we must learn to make the same distinction for the problem of the inductive relation from facts to theories. (Reichenbach 1938: 36–37)

The standard interpretation of this statement marks contextual values, which may have contributed to the discovery of a theory, as irrelevant for *justifying* the acceptance of a theory, and for assessing how evidence

bears on theory—the relation that is crucial for the objectivity of science. Contextual values are restricted to a matter of individual psychology that may influence the discovery, development and proliferation of a scientific theory, but not its epistemic status.

This distinction played a crucial role in post-World War II philosophy of science. It presupposes, however, a clear-cut distinction between cognitive values on the one hand and contextual values on the other. While this may be *prima facie* plausible for disciplines such as physics, there is an abundance of contextual values in the social sciences, for instance, in the conceptualization and measurement of a nation's wealth, or in different ways to measure the inflation rate (cf. Dupré 2007; Reiss 2008). More generally, three major lines of criticism can be identified.

First, Helen Longino (1996) has argued that traditional cognitive values such as consistency, simplicity, breadth of scope and fruitfulness are not purely cognitive or epistemic after all, and that their use imports political and social values into contexts of scientific judgment. According to her, the use of cognitive values in scientific judgments is not always, not even normally, politically neutral. She proposes to juxtapose these values with feminist values such as novelty, ontological heterogeneity, mutuality of interaction, applicability to human needs and diffusion of power, and argues that the use of the traditional value instead of its alternative (e.g., simplicity instead of ontological heterogeneity) can lead to biases and adverse research results. Longino's argument here is different from the one discussed in [section 3.1](#). It casts the very distinction between cognitive and contextual values into doubt.

The second argument against the possibility of value-free science is semantic and attacks the neutrality of scientific theories: fact and value are frequently entangled because of the use of so-called “thick” ethical concepts in science (Putnam 2002)—i.e., ethical concepts that have mixed descriptive and normative content. For example, a description such as “dangerous technology” involves a value judgment about the technology and the risks it implies, but it also has a descriptive content: it is uncertain and hard to predict whether using that technology will really trigger those risks. If the use of such terms, where facts and values are inextricably entangled, is inevitable in scientific reasoning, it is impossible to describe hypotheses and results in a value-free manner, undermining the value-neutrality thesis.

Indeed, John Dupré has argued that thick ethical terms are ineliminable from science, at least certain parts of it (Dupré 2007). Dupré's point is essentially that scientific hypotheses and results concern us because they are relevant to human interests, and thus they will necessarily be couched in a language that uses thick ethical terms. While it will often be possible to translate ethically thick descriptions into neutral ones, the translation cannot be made without losses, and these losses obtain precisely because human interests are involved (see [section 6.2](#) for a case study from social science). According to Dupré, then, many scientific statements are value-free only because their truth or falsity does not matter to us:

Whether electrons have a positive or a negative charge and whether there is a black hole in the middle of our galaxy are questions of absolutely no immediate importance to us. The only human interests they touch (and these they may indeed touch deeply) are cognitive ones, and so the only values that they implicate are cognitive values. (2007: 31)

A third challenge to the VNT, and perhaps the most influential one, was raised first by Richard Rudner in his influential article “The Scientist Qua Scientist Makes Value Judgments” (Rudner 1953). Rudner disputes the core of the VNT and the context of discovery/justification distinction: the idea that the acceptance of a scientific theory can in principle be value-free. First, Rudner argues that

no analysis of what constitutes the method of science would be satisfactory unless it comprised some assertion to the effect that the scientist as scientist **accepts or rejects hypotheses**. (1953: 2)

This assumption stems from industrial quality control and other application-oriented research. In such contexts, it is often necessary to accept or to reject a hypothesis (e.g., the efficacy of a drug) in order to make effective decisions.

Second, he notes that no scientific hypothesis is ever confirmed beyond reasonable doubt—some probability of error always remains. When we accept or reject a hypothesis, there is always a chance that our decision is mistaken. Hence, our decision is also “a function of the *importance*, in the typically ethical sense, of making a mistake in accepting or rejecting a hypothesis” (1953: 2): we are balancing the seriousness of two possible errors (erroneous acceptance/rejection of the hypothesis) against each other. This corresponds to type I and type II error in statistical inference.

The decision to accept or reject a hypothesis involves a value judgment (at least implicitly) because scientists have to judge which of the consequences of an erroneous decision they deem more palatable: (1) some individuals die of the side effects of a drug erroneously judged to be safe; or (2) other individuals die of a condition because they did not have access to a treatment that was erroneously judged to be unsafe. Hence, ethical judgments and contextual values necessarily enter the scientist’s core activity of accepting and rejecting hypotheses, and the VNT stands refuted. Closely related arguments can be found in Churchman (1948) and Braithwaite (1953). Hempel (1965: 91–92) gives a modified account of Rudner’s argument by distinguishing between judgments of *confirmation*, which are free of contextual values, and judgments of *acceptance*. Since even strongly confirming evidence cannot fully prove a universal scientific law, we have to live with a residual “inductive risk” in inferring that law. Contextual values influence scientific methods by determining the acceptable amount of inductive risk (see also Douglas 2000).

But how general are Rudner’s objections? Apparently, his result holds true of applied science, but not necessarily of fundamental research. For the latter domain, two major lines of rebuttals have been proposed. First, Richard Jeffrey (1956) notes that lawlike hypotheses in theoretical science (e.g., the gravitational law in Newtonian mechanics) are characterized by their general scope and not confined to a particular application. Obviously, a scientist cannot fine-tune her decisions to their possible consequences in a wide variety of different contexts. So she should just refrain from the essentially pragmatic decision to accept or reject hypotheses. By restricting scientific reasoning to gathering and interpreting evidence, possibly supplemented by assessing the probability of a hypothesis, Jeffrey tries to save the VNT in fundamental scientific research, and the objectivity of scientific reasoning.

Second, Isaac Levi (1960) observes that scientists commit themselves to certain standards of inference when they become a member of the profession. This may, for example, lead to the statistical rejection of a hypothesis when the observed significance level is smaller than 5%. These community standards may eliminate any room for contextual ethical judgment on behalf of the scientist: they determine when she should accept a hypothesis as established. Value judgments may be implicit in how a scientific community sets standards of inference (compare [section 5.1](#)), but not in the daily work of an *individual* scientist (cf. Wilholt 2013).

Both defenses of the VNT focus on the impact of values in theory choice, either by denying that scientists actually choose theories (Jeffrey), or by referring to community standards and restricting the VNT to the individual scientist (Levi). Douglas (2000: 563–565) points out, however, that the “acceptance” of scientific theories is only one of several places for values to enter scientific reasoning, albeit an especially prominent and explicit one. Many decisions in the process of scientific inquiry may conceal implicit value judgments: the design of an experiment, the methodology for conducting it, the characterization of the data, the choice of a statistical method for processing and analyzing data, the interpretational process findings, etc. None of these methodological decisions could be made without consideration of the possible consequences that could occur. Douglas gives, as a case study, a series of experiments where carcinogenic effects of dioxin exposure on rats were probed. Contextual values such as safety and risk aversion affected the conducted research at various stages: first, in the classification of pathological samples as benign or cancerous (over which a lot of expert disagreement occurred), second, in the extrapolation from the high-dose experimental conditions to the more realistic low-dose conditions. In both cases, the choice of a conservative classification or model had to be weighed against the adverse consequences for society that could result from underestimating the risks (see also Biddle 2013).

These diagnoses cast a gloomy light on attempts to divide scientific labor between gathering evidence and determining the degree of confirmation (value-free) on the one hand and accepting scientific theories (value-laden) on the other. The entire process of conceptualizing, gathering and interpreting evidence is so entangled

with contextual values that no neat division, as Jeffrey envisions, will work outside the narrow realm of statistical inference—and even there, doubts may be raised ([see section 4.2](#)).

Philip Kitcher (2011a: 31–40; see also Kitcher 2011b) gives an alternative argument, based on his idea of “significant truths”. There are simply too many truths that are of no interest whatsoever, such as the total number of offside positions in a low-level football competition. Science, then, doesn’t aim at truth *simpliciter* but rather at something more narrow: truth worth pursuing from the point of view of our cognitive, practical and social goals. Any truth that is worth pursuing in this sense is what he calls a “significant truth”. Clearly, it is value judgments that help us decide whether or not any given truth is significant.

Kitcher goes on to observing that the process of scientific investigation cannot neatly be divided into a stage in which the research question is chosen, one in which the evidence is gathered and one in which a judgment about the question is made on the basis of the evidence. Rather, the sequence is multiply iterated, and at each stage, the researcher has to decide whether previous results warrant pursuit of the current line of research, or whether she should switch to another avenue. Such choices are laden with contextual values.

Values in science also interact, according to Kitcher, in a non-trivial way. Assume we endorse predictive accuracy as an important goal of science. However, there may not be a convincing strategy to reach this goal in some domain of science, for instance because that domain is characterized by strong non-linear dependencies. In this case, predictive accuracy might have to yield to achieving other values, such as consistency with theories in neighbor domains. Conversely, changing social goals lead to re-evaluations of scientific knowledge and research methods.

Science, then, cannot be value-free because no scientist ever works exclusively in the supposedly value-free zone of assessing and accepting hypotheses. Evidence is gathered and hypotheses are assessed and accepted in the light of their potential for application and fruitful research avenues. Both cognitive and contextual value judgments guide these choices and are themselves influenced by their results.

### 3.3 Science, Policy and the Value-Free Ideal

The discussion so far has focused on the VNT, the practical attainability of the VFI, but little has been said about whether a value-free science is desirable in the first place. This subsection discusses this topic with special attention to informing and advising public policy from a scientific perspective. While the VFI, and many arguments for and against it, can be applied to science as a whole, the interface of science and public policy is the place where the intrusion of values into science is especially salient, and where it is surrounded by the greatest controversy. In the 2009 “Climategate” affair, leaked emails from climate scientists raised suspicions that they were pursuing a particular socio-political agenda that affected their research in an improper way. Later inquiries and reports absolved them from charges of misconduct, but the suspicions alone did much to damage the authority of science in the public arena.

Indeed, many debates at the interface of science and public policy are characterized by disagreements on propositions that combine a factual basis with specific goals and values. Take, for instance, the view that growing transgenic crops carries too much risk in terms of biosecurity, or addressing global warming by phasing out fossil energies immediately. The critical question in such debates is whether there are theses  $T$  such that one side in the debate endorses  $T$ , the other side rejects it, the evidence is shared, and both sides have good reasons for their respective positions.

According to the VFI, scientists should uncover an epistemic, value-free basis for resolving such disagreements and restrict the dissent to the realm of value judgments. Even if the VNT should turn out to be untenable, and a strict separation to be impossible, the VFI may have an important function for *guiding* scientific research and for minimizing the impact of values on an objective science. In the philosophy of science, one camp of scholars defends the VFI as a necessary antidote to individual and institutional interests, such as Hugh Lacey (1999, 2002), Ernan McMullin (1982) and Sandra Mitchell (2004), while others adopt a critical attitude, such as Helen

Longino (1990, 1996), Philip Kitcher (2011a) and Heather Douglas (2009). These criticisms we discuss mainly refer to the desirability or the conceptual (un)clarity of the VFI.

First, it has been argued that the VFI is not desirable at all. Feminist philosophers (e.g., Harding 1991; Okruhlik 1994; Lloyd 2005) have argued that science often carries a heavy androcentric values, for instance in biological theories about sex, gender and rape. The charge against these values is not so much that they are contextual rather than cognitive, but that they are unjustified. Moreover, if scientists did follow the VFI rigidly, policy-makers would pay even less attention to them, with a detrimental effect on the decisions they take (Cranor 1993). Given these shortcomings, the VFI has to be rethought if it is supposed to play a useful role for guiding scientific research and leading to better policy decisions. [Section 4.3](#) and [section 5.2](#) elaborate on this line of criticism in the context of scientific community practices, and a science in the service of society.

Second, the autonomy of science often fails in practice due to the presence of external stakeholders, such as funding agencies and industry lobbies. To save the epistemic authority of science, Douglas (2009: 7–8) proposes to detach it from its autonomy by reformulating the VFI and distinguishing between **direct and indirect roles of values in science**. Contextual values may legitimately affect the assessment of evidence by indicating the appropriate standard of evidence, the representation of complex processes, the severity of consequences of a decision, the interpretation of noisy datasets, and so on (see also Winsberg 2012). This concerns, above all, policy-related disciplines such as climate science or economics that routinely perform scientific risk analyses for real-world problems (cf. also Shrader-Frechette 1991). Values should, however, not be “reasons in themselves”, that is, evidence or defeaters for evidence (direct role, illegitimate) and as “helping to decide what should count as a *sufficient* reason for a choice” (indirect role, legitimate). This prohibition for values to replace or dismiss scientific evidence is called **detached objectivity** by Douglas, but it is complemented by various other aspects that relate to a reflective balancing of various perspectives and the procedural, social aspects of science (2009: ch. 6).

That said, Douglas’ proposal is not very concrete when it comes to implementation, e.g., regarding the way diverse values should be balanced. Compromising in the middle cannot be the solution (Weber 1917 [1949]). First, no standpoint is, just in virtue of being in the middle, evidentially supported vis-à-vis more extreme positions. Second, these middle positions are also, from a practical point of view, the least functional when it comes to advising policy-makers.

Moreover, the distinction between direct and indirect roles of values in science may not be sufficiently clear-cut to police the legitimate use of values in science, and to draw the necessary borderlines. Assume that a scientist considers, for whatever reason, the consequences of erroneously accepting hypothesis  $H$  undesirable. Therefore he uses a statistical model whose results are likely to favor  $\neg H$  over  $H$ . Is this a matter of reasonable conservativeness? Or doesn’t it amount to reasoning to a foregone conclusion, and to treating values as evidence (cf. Elliott 2011: 320–321)?

The most recent literature on values and evidence in science presents us with a broad spectrum of opinions. Steele (2012) and Winsberg (2012) agree that probabilistic assessments of uncertainty involve contextual value judgments. While Steele defends this point by analyzing the role of scientists as policy advisors, Winsberg points to the influence of contextual values in the selection and representation of physical processes in climate modeling. Betz (2013) argues, by contrast, that scientists can largely avoid making contextual value judgments if they carefully express the uncertainty involved with their evidential judgments, e.g., by using a scale ranging from purely qualitative evidence (such as expert judgment) to precise probabilistic assessments. The issue of value judgments at earlier stages of inquiry is not addressed by this proposal; however, disentangling evidential judgments and judgments involving contextual values at the stage of theory assessment may be a good thing in itself.

Thus, should we or should we not worried about values in scientific reasoning? While the interplay of values and evidential considerations need not be pernicious, it is unclear why it *adds* to the success or the authority of science. How are we going to ensure that the permissive attitude towards values in setting evidential standards etc. is not abused? In the absence of a general theory about which contextual values are beneficial and which are pernicious, the VFI might as well be as a first-order approximation to a sound, transparent and objective science.

## 4. Objectivity as Freedom from Personal Biases

This section deals with scientific objectivity as a form of intersubjectivity—as freedom from personal biases. According to this view, science is objective to the extent that personal biases are absent from scientific reasoning, or that they can be eliminated in a social process. Perhaps all science is necessarily perspectival. Perhaps we cannot sensibly draw scientific inferences without a host of background assumptions, which may include assumptions about values. Perhaps all scientists are biased in some way. But objective scientific results do not, or so the argument goes, depend on researchers' personal preferences or experiences—they are the result of a process where individual biases are gradually filtered out and replaced by agreed upon evidence. That, among other things, is what distinguishes science from the arts and other human activities, and scientific knowledge from a fact-independent social construction (e.g., Haack 2003).

Paradigmatic ways to achieve objectivity in this sense are measurement and quantification. What has been measured and quantified has been verified relative to a standard. The truth, say, that the Eiffel Tower is 324 meters tall is relative to a standard unit and conventions about how to use certain instruments, so it is neither aperspectival nor free from assumptions, but it is independent of the person making the measurement.

We will begin with a discussion of objectivity, so conceived, in measurement, discuss the ideal of “mechanical objectivity” and then investigate to what extent freedom from personal biases can be implemented in statistical evidence and inductive inference—arguably the core of scientific reasoning, especially in quantitatively oriented sciences. Finally, we discuss Feyerabend's radical criticism of a rational scientific method that can be mechanically applied, and his defense of the epistemic and social benefits of personal “bias” and idiosyncrasy.

### 4.1 Measurement and Quantification

Measurement is often thought to epitomize scientific objectivity, most famously captured in Lord Kelvin's dictum

when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of *science*, whatever the matter may be. (Kelvin 1883, 73)

Measurement can certainly achieve some independence of perspective. Yesterday's weather in Durham UK may have been “really hot” to the average North Eastern Brit and “very cold” to the average Mexican, but they'll both accept that it was 21°C. Clearly, however, measurement does not result in a “view from nowhere”, nor are typical measurement results free from presuppositions. Measurement instruments interact with the environment, and so results will always be a product of both the properties of the environment we aim to measure as well as the properties of the instrument. Instruments, thus, provide a perspectival view on the world (cf. Giere 2006).

Moreover, making sense of measurement results requires interpretation. Consider temperature measurement. Thermometers function by relating an unobservable quantity, temperature, to an observable quantity, expansion (or length) of a fluid or gas in a glass tube; that is, thermometers measure temperature by assuming that length is a function of temperature:  $\text{length} = f(\text{temperature})$ . The function  $f$  is not known *a priori*, and it cannot be tested either (because it could in principle only be tested using a *veridical* thermometer, and the veridicality of the thermometer is just what is at stake here). Making a specific assumption, for instance that  $f$  is linear, solves that problem by fiat. But this “solution” does not take us very far because different thermometric substances (e.g., mercury, air or water) yield different results for the points intermediate between the two fixed points 0°C and 100°C, and so they can't all expand linearly.

According to Hasok Chang's account of early thermometry (Chang 2004), the problem was eventually solved by using a “principle of minimalist overdetermination”, the goal of which was to find a reliable thermometer while making as few substantial assumptions (e.g., about the form for  $f$ ) as possible. It was argued that if a thermometer was to be reliable, different tokens of the same thermometer type should agree with each other, and the results of air thermometers agreed the most. “Minimal” doesn't mean zero, however, and indeed this

procedure makes an important presupposition (in this case a metaphysical assumption about the one-valuedness of a physical quantity). Moreover, the procedure yielded at best a reliable instrument, not necessarily one that was best at tracking the uniquely real temperature (if there is such a thing).

What Chang argues about early thermometry is true of measurements more generally: they are always made against a backdrop of metaphysical presuppositions, theoretical expectations and other kinds of belief. Whether or not any given procedure is regarded as adequate depends to a large extent on the purposes pursued by the individual scientist or group of scientists making the measurements. Especially in the social sciences, this often means that measurement procedures are laden with normative assumptions, i.e., values.

Julian Reiss (2008, 2013) has argued that economic indicators such as consumer price inflation, gross domestic product and the unemployment rate are value-laden in this sense. Consumer-price indices, for instance, assume that if a consumer prefers a bundle  $x$  over an alternative  $y$ , then  $x$  is better for her than  $y$ , which is as ethically charged as it is controversial. National income measures assume that nations that exchange a larger share of goods and services on markets are richer than nations where the same goods and services are provided by the government or within households, which too is ethically charged and controversial.

While not free of assumptions and values, the goal of many measurement procedures remains to reduce the influence of personal biases and idiosyncrasies. The Nixon administration, famously, indexed social security payments to the consumer-price index in order to eliminate the dependence of security recipients on the flimsiest of party politics: to make increases automatic instead of a result of political negotiations (Nixon 1969). Lorraine Daston and Peter Galison refer to this as **mechanical objectivity**. They write:

Finally, we come to the full-fledged establishment of mechanical objectivity as the ideal of scientific representation. What we find is that the image, as standard bearer of objectivity is tied to a relentless search to replace individual volition and discretion in depiction by the invariable routines of mechanical reproduction. (Daston and Galison 1992: 98)

Mechanical objectivity reduces the importance of human contributions to scientific results to a minimum, and therefore enables science to proceed on a large scale where bonds of trust between individuals can no longer hold (Daston 1992). Trust in mechanical procedures thus replaces trust in individual scientists.

In his book *Trust in Numbers*, Theodore Porter pursues this line of thought in great detail. In particular, on the basis of case studies involving British actuaries in the mid-nineteenth century, of French state engineers throughout the century, and of the US Army Corps of Engineers from 1920 to 1960, he argues for two causal claims. First, measurement instruments and quantitative procedures originate in commercial and administrative needs and affect the ways in which the natural and social sciences are practiced, not the other way around. The mushrooming of instruments such as chemical balances, barometers, chronometers was largely a result of social pressures and the demands of democratic societies. Administering large territories or controlling diverse people and processes is not always possible on the basis of personal trust and thus “objective procedures” (which do not require trust in persons) took the place of “subjective judgments” (which do). Second, he argues that quantification is a technology of distrust and weakness, and not of strength. It is weak administrators who do not have the social status, political support or professional solidarity to defend their experts’ judgments. They therefore subject decisions to public scrutiny, which means that they must be made in a publicly accessible form.

This is the situation in which scientists who work in areas where the science/policy boundary is fluid find themselves:

The National Academy of Sciences has accepted the principle that scientists should declare their conflicts of interest and financial holdings before offering policy advice, or even information to the government. And while police inspections of notebooks remain exceptional, the personal and financial interests of scientists and engineers are often considered material, especially in legal and regulatory contexts.

Strategies of impersonality must be understood partly as defenses against such suspicions [...]. Objectivity means knowledge that does not depend too much on the particular individuals who author it. (Porter 1995: 229)

Measurement and quantification help to reduce the influence of personal biases and idiosyncrasies and they reduce the need to trust the scientist or government official, but often at a cost. Standardizing scientific procedures becomes difficult when their subject matters are not homogeneous, and few domains outside fundamental physics are. Attempts to quantify procedures for treatment and policy decisions that we find in evidence-based practices are currently transferred to a variety of sciences such as medicine, nursing, psychology, education and social policy. However, they often lack a certain degree of responsiveness to the peculiarities of their subjects and the local conditions to which they are applied (see also [section 5.3](#)).

Moreover, the measurement and quantification of characteristics of scientific interest is only half of the story. We also want to describe relations between the quantities and make inferences using statistical analysis. Statistics thus helps to quantify further aspects of scientific work. We will now examine whether or not statistical analysis can proceed in a way free from personal biases and idiosyncrasies—for more detail, see the entry on [philosophy of statistics](#).

## 4.2 Statistical Evidence

The appraisal of scientific evidence is traditionally regarded as a domain of scientific reasoning where the ideal of scientific objectivity has strong normative force, and where it is also well-entrenched in scientific practice. Episodes such as Galilei’s observations of the Jupiter moons, Lavoisier’s calcination experiments, and Eddington’s observation of the 1919 eclipse are found in all philosophy of science textbooks because they exemplify how evidence can be persuasive and compelling to scientists with different backgrounds. The crucial question is therefore: can we identify an “objective” concept of scientific evidence that is independent of the personal biases of the experimenter and interpreter?

Inferential statistics—the field that investigates the validity of inferences from data to theory—tries to answer this question. It is extremely influential in modern science, pervading experimental research as well as the assessment and acceptance of our most fundamental theories. For instance, a statistical argument helped to establish the recent discovery of the Higgs Boson. We now compare the main theories of statistical evidence with respect to the objectivity of the claims they produce. They mainly differ with respect to the role of an explicitly subjective interpretation of probability.

### 4.2.1 Bayesian Inference

Bayesian inference quantifies scientific evidence by means of probabilities that are interpreted as a scientist’s subjective degrees of belief. The Bayesian thus leaves behind Carnap’s (1950) idea that probability is determined by a logical relation between sentences. For example, the prior degree of belief in hypothesis  $H$ , written  $p(H)$ , can in principle take any value in the interval  $[0, 1]$ . Simultaneously held degrees of belief in different hypotheses are, however, constrained by the laws of probability. After learning evidence  $E$ , the degree of belief in  $H$  is changed from its prior probability  $p(H)$  to the conditional degree of belief  $p(H | E)$ , commonly called the posterior probability of  $H$ . Both quantities can be related to each other by means of [Bayes’ Theorem](#).

These days, the Bayesian approach is extremely influential in philosophy and rapidly gaining ground across all scientific disciplines. For quantifying evidence for a hypothesis, Bayesian statisticians almost uniformly use the **Bayes factor**, that is, the ratio of prior to posterior odds in favor of a hypothesis. The Bayes factor in favor of hypothesis  $H$  against its negation  $\neg H$  in the light of evidence  $E$  can be written as

$$(3) \quad BF(E) := \frac{p(H | E)}{p(\neg H | E)} \cdot \frac{p(\neg H)}{p(H)} = \frac{p(E | H)}{p(E | \neg H)},$$

or in other words, as the likelihood ratio between  $H$  and  $\neg H$ . The Bayes factor reduces to the likelihoodist conception of evidence (Royall 1997) for the case of two competing point hypotheses. For further discussion of Bayesian measures of evidence, see Good (1950), Sprenger and Hartmann (2019: ch. 1) and the entry on [confirmation and evidential support](#).

Unsurprisingly, the idea to measure scientific evidence in terms of subjective probability has met resistance. For example, the statistician Ronald A. Fisher (1935: 6–7) has argued that measuring psychological tendencies cannot be relevant for scientific inquiry and sustain claims to objectivity. Indeed, how should scientific objectivity square with subjective degree of belief? Bayesians have responded to this challenge in various ways:

- Howson (2000) and Howson and Urbach (2006) consider the objection misplaced. In the same way that deductive logic does not judge the correctness of the premises but just advises you what to infer from them, **Bayesian inductive logic** provides rational rules for representing uncertainty and making inductive inferences. Choosing the premises (e.g., the prior distributions) “objectively” falls outside the scope of Bayesian analysis.
- **Convergence or merging-of-opinion theorems** guarantee that under certain circumstances, agents with very different initial attitudes who observe the same evidence will obtain similar posterior degrees of belief in the long run. However, they are asymptotic results without direct implications for inference with real-life datasets (see also Earman 1992: ch. 6). In such cases, the choice of the prior matters, and it may be beset with idiosyncratic bias and manifest social values.
- Adopting a more modest stance, Sprenger (2018) accepts that Bayesian inference does not achieve the goal of objectivity in the sense of intersubjective agreement (concordant objectivity), or being free of personal values, bias and subjective judgment. However, he argues that competing schools of inference such as frequentist inference face this problem to the same degree, perhaps even worse. Moreover, some features of Bayesian inference (e.g., the transparency about prior assumptions) fit recent, socially oriented conceptions of objectivity that we discuss in [section 5](#).

A radical Bayesian solution to the problem of personal bias is to adopt a principle that radically constrains an agent’s rational degrees of belief, such as the **Principle of Maximum Entropy** (MaxEnt—Jaynes 1968; Williamson 2010). According to MaxEnt, degrees of belief must be probabilistic and in sync with empirical constraints, but conditional on these constraints, they must be equivocal, that is, as middling as possible. This latter constraint amounts to maximizing the entropy of the probability distribution in question. The MaxEnt approach eliminates various sources of subjective bias at the expense of narrowing down the range of rational degrees of belief. An alternative objective Bayesian solution consists in so-called “**objective priors**”: prior probabilities that do not represent an agent’s factual attitudes, but are determined by principles of symmetry, mathematical convenience or maximizing the influence of the data on the posterior (e.g., Jeffreys 1939 [1980]; Bernardo 2012).

Thus, Bayesian inference, which analyzes statistical evidence from the vantage point of rational belief, provides only a partial answer to securing scientific objectivity from personal idiosyncrasy.

#### 4.2.2 Frequentist Inference

The frequentist conception of evidence is based on the idea of the **statistical test of a hypothesis**. Under the influence of the statisticians Jerzy Neyman and Egon Pearson, tests were often regarded as rational decision procedures that minimize the relative frequency of wrong decisions in a hypothetical series of repetitions of a test (hence the name “frequentism”). Rudner’s argument in [section 3.2](#) has pointed out the limits of this conception of hypothesis tests: the choice of thresholds for acceptance and rejection (i.e., the acceptable type I and II error rates) may reflect contextual value judgments and personal bias. Moreover, the losses associated with erroneously accepting or rejecting that hypothesis depend on the context of application which may be unbeknownst to the experimenter.

Alternatively, scientists can restrict themselves to a purely evidential interpretation of hypothesis tests and leave decisions to policy-makers and regulatory agencies. The statistician and biologist R.A. Fisher (1935, 1956) proposed what later became the orthodox quantification of evidence in frequentist statistics. Suppose a “null” or default hypothesis  $H_0$  denotes that an intervention has zero effect. If the observed data are “extreme” under  $H_0$ —i.e., if it was highly likely to observe a result that agrees better with  $H_0$ —the data provide evidence against the null hypothesis and for the efficacy of the intervention. The epistemological rationale is connected to the idea of severe testing (Mayo 1996): if the intervention were ineffective, we would, in all likelihood, have found data that agree better with the null hypothesis. The strength of evidence against  $H_0$  is equal to the ***p*-value**: the lower it is, the stronger evidence  $E$  speaks against the null hypothesis  $H_0$ .

Unlike Bayes factors, this concept of statistical evidence does not depend on personal degrees of belief. However, this does not necessarily mean that *p*-values are more objective. First, *p*-values are usually classified as “non-significant” ( $p > .05$ ), “significant” ( $p < .05$ ), “highly significant”, and so on. Not only that these thresholds and labels are largely arbitrary, they also promote **publication bias**: non-significant findings are often classified as “failed studies” (i.e., the efficacy of the intervention could not be shown), rarely published and end up in the proverbial “file drawer”. Much valuable research is suppressed. Conversely, significant findings may often occur when the null hypothesis is actually true, especially when researchers have been “hunting for significance”. In fact, researchers have an incentive to keep their *p*-values low: the stronger the evidence, the more convincing the narrative, the greater the impact—and the higher the chance for a good publication and career-relevant rewards. Moving the goalpost by “*p*-hacking” outcomes—for example by eliminating outliers, selective reporting or restricting the analysis to a subgroup—evidently biases the research results and compromises the objectivity of experimental research.

In particular, such **questionable research practices (QRP)** increase the type I error rate, which measures the rate at which false hypotheses are accepted, substantially over its nominal 5% level and contribute to publication bias (Bakker et al. 2012). Ioannidis (2005) concludes that “most published research findings are false”—they are the combined result of a low base rate of effective causal interventions, the file drawer effect and the widespread presence of questionable research practices. The frequentist logic of hypothesis testing aggravates the problem because it provides a framework where all these biases can easily enter (Ziliak and McCloskey 2008; Sprenger 2016). These radical conclusions are also confirmed by empirical findings: in many disciplines researchers fail to replicate findings by other scientific teams. See [section 5.1](#) for more detail.

Summing up our findings, neither of the two major frameworks of statistical inference manages to eliminate all sources of personal bias and idiosyncrasy. The Bayesian considers subjective assumptions to be an irreducible part of scientific reasoning and sees no harm in making them explicit. The frequentist conception of evidence based on *p*-values avoids these explicitly subjective elements, but at the price of a misleading impression of objectivity and frequent abuse in practice. A defense of frequentist inference should, in our opinion, stress that the relatively rigid rules for interpreting statistical evidence facilitate communication and assessment of research results in the scientific community—something that is harder to achieve for a Bayesian. We now turn from specific methods for stating and interpreting evidence to a radical criticism of the idea that there is a rational scientific method.

### 4.3 Feyerabend: The Tyranny of the Rational Method

In his writings of the 1970s, [Paul Feyerabend](#) launched a profound attack on the rationality and objectivity of scientific method. His position is exceptional in the philosophical literature since traditionally, the threat for objective and successful science is located in contextual rather than epistemic values. Feyerabend turns this view upside down: it is the “tyranny” of rational method, and the emphasis on epistemic rather than contextual values that prevents us from having a science in the service of society. Moreover, he welcomes a diversity of different personal, also idiosyncratic perspectives, thus denying the idea that freedom from personal “bias” is epistemically and socially beneficial.

The starting point of Feyerabend’s criticism of rational method is the thesis that strict epistemic rules such as those expressed by the VFI only suppress an open exchange of ideas, extinguish scientific creativity and prevent

a free and truly democratic science. In his classic “Against Method” (1975: chs. 8–13), Feyerabend elaborates on this criticism by examining a famous episode in the history of science. When the Catholic Church objected to Galilean mechanics, it had the better arguments by the standards of seventeenth-century science. Their conservatism in their position was scientifically backed: Galilei’s telescopes were unreliable for celestial observations, and many well-established phenomena (no fixed star parallax, invariance of laws of motion) could not yet be explained in the heliocentric system. With hindsight, Galilei managed to achieve groundbreaking scientific progress just because he deliberately violated rules of scientific reasoning. Hence Feyerabend’s dictum “Anything goes”: no methodology whatsoever is able to capture the creative and often irrational ways by which science deepens our understanding of the world. Good scientific reasoning cannot be captured by rational method, as Carnap, Hempel and Popper postulated.

The drawbacks of an objective, value-free and method-bound view on science and scientific method are not only epistemic. Such a view narrows down our perspective and makes us less free, open-minded, creative, and ultimately, less human in our thinking (Feyerabend 1975: 154). It is therefore neither possible nor desirable to have an objective, value-free science (cf. Feyerabend 1978: 78–79). As a consequence, Feyerabend sees traditional forms of inquiry about our world (e.g., Chinese medicine) on a par with their Western competitors. He denounces appeals to “objective” standards as rhetorical tools for bolstering the epistemic authority of a small intellectual elite (=Western scientists), and as barely disguised statements of preference for one’s own worldview:

there is hardly any difference between the members of a “primitive” tribe who defend their laws because they are the laws of the gods [...] and a rationalist who appeals to “objective” standards, except that the former know what they are doing while the latter does not. (1978: 82)

In particular, when discussing other traditions, we often project our own worldview and value judgments into them instead of making an impartial comparison (1978: 80–83). There is no purely rational justification for dismissing other perspectives in favor of the Western scientific worldview—the insistence on our Western approach may be as justified as insisting on absolute space and time after the Theory of Relativity.

The Galilei example also illustrates that personal perspective and idiosyncratic “bias” need not be bad for science. Feyerabend argues further that scientific research is accountable to society and should be kept in check by democratic institutions, and laymen in particular. Their particular perspectives can help to determine the funding agenda and to set ethical standards for scientific inquiry, but also be useful for traditionally value-free tasks such as choosing an appropriate research method and assessing scientific evidence. Feyerabend’s writings on this issue were much influenced by witnessing the Civil Rights Movement in the U.S. and the increasing emancipation of minorities, such as Blacks, Asians and Hispanics.

All this is not meant to say that truth loses its function as a normative concept, nor that all scientific claims are equally acceptable. Rather, Feyerabend advocates an **epistemic pluralism** that accepts diverse approaches to acquiring knowledge. Rather than defending a narrow and misleading ideal of objectivity, science should respect the diversity of values and traditions that drive our inquiries about the world (1978: 106–107). This would put science back into the role it had during the scientific revolution or the Enlightenment: as a liberating force that fought intellectual and political oppression by the sovereign, the nobility or the clergy. Objections to this view are discussed at the end of [section 5.2](#).

## 5. Objectivity as a Feature of Scientific Communities and Their Practices

This section addresses various accounts that regard scientific objectivity essentially as a function of social practices in science and the social organization of the scientific community. All these accounts reject the characterization of scientific objectivity as a function of correspondence between theories and the world, as a feature of individual reasoning practices, or as pertaining to individual studies and experiments (see also Douglas 2011). Instead, they evaluate the objectivity of a *collective* of studies, as well as the methods and

community practices that structure and guide scientific research. More precisely, they adopt a meta-analytic perspective for assessing the reliability of scientific results (section 5.1), and they construct objectivity from a feminist perspective: as an open interchange of mutual criticism, or as being anchored in the “situatedness” of our scientific practices and the knowledge we gain ([section 5.2](#)).

## 5.1 Reproducibility and the Meta-Analytic Perspective

The collectivist perspective is especially useful when an entire discipline enters a stage of crisis: its members become convinced that a significant proportion of findings are not trustworthy. A contemporary example of such a situation is the **replication crisis**, which was briefly mentioned in the previous section and concerns the *reproducibility* of scientific knowledge claims in a variety of different fields (most prominently: psychology, biology, medicine). Large-scale replication projects have noticed that many findings which we considered as an integral part of scientific knowledge failed to replicate in settings that were designed to mimic the original experiment as closely as possible (e.g., Open Science Collaboration 2015). Successful attempts at replicating an experimental result have long been argued to provide evidence of freedom from particular kinds of artefacts and thus the trustworthiness of the result. Compare the [entry on experiment in physics](#). Likewise, failure to replicate indicates that either the original finding, the result of the replication attempt, or both, are biased—though see John Norton’s (ms., ch. 3—see Other Internet Resources) arguments that the evidential value of (failed) replications crucially depends on researchers’ material background assumptions.

When replication failures in a discipline are particularly significant, one may conclude that the published literature lacks objectivity—at a minimum the discipline fails to inspire trust that its findings are more than artefacts of the researchers’ efforts. Conversely, when observed effects can be replicated in follow-up experiments, a kind of objectivity is reached that goes beyond the ideas of freedom from personal bias, mechanical objectivity, and subject-independent measurement, discussed in [section 4.1](#).

Freese and Peterson (2018) call this idea **statistical objectivity**. It grounds in the view that even the most scrupulous and diligent researchers cannot achieve full objectivity all by themselves. The term “objectivity” instead applies to a collection or population of studies, with **meta-analysis** (a formal method for aggregating the results from ranges of studies) as the “apex of objectivity” (Freese and Peterson 2018, 304; see also Stegenga 2011, 2018). In particular, aggregating studies from different researchers may provide evidence of systematic bias and questionable research practices (QRP) in the published literature. This diagnostic function of meta-analysis for detecting violations of objectivity is enhanced by statistical techniques such as the funnel plot and the *p*-curve (Simonsohn et al. 2014).

Apart from this epistemic dimension, research on statistical objectivity also has an activist dimension: methodologists urge researchers to make publicly available essential parts of their research before the data analysis starts, and to make their methods and data sources more transparent. For example, it is conjectured that the replicability (and thus objectivity) of science will increase by making all data available online, by preregistering experiments, and by using the registered reports model for journal articles (i.e., the journal decides on publication before data collection on the basis of the significance of the proposed research as well as the experimental design). The idea is that transparency about the data set and the experimental design will make it easier to stage a replication of an experiment and to assess its methodological quality. Moreover, publicly committing to a data analysis plan beforehand will lower the rate of QRPs and of attempts to *accommodate* data to hypotheses rather than making proper predictions.

All in all, statistical objectivity moves the discussion of objectivity to the level of population of studies. There, it takes up and modifies several conceptions of objectivity that we have seen before: most prominently, freedom of subjective bias, which is replaced with collective bias and pernicious conventions, and the subject-independent measurement of a physical quantity, which is replaced by reproducibility of effects.

## 5.2 Feminist and Standpoint Epistemology

Traditional notions of objectivity as faithfulness to facts or freedom of contextual values have also been challenged from a feminist perspective. These critiques can be grouped in three major research programs: feminist epistemology, feminist standpoint theory and feminist postmodernism (Crasnow 2013). The program of **feminist epistemology** explores the impact of sex and gender on the production of scientific knowledge. More precisely, feminist epistemology highlights the epistemic risks resulting from the systematic exclusion of women from the ranks of scientists, and the neglect of women as objects of study. Prominent case studies are the neglect of female orgasm in biology, testing medical drugs on male participants only, focusing on male specimen when studying the social behavior of primates, and explaining human mating patterns by means of imaginary neolithic societies (e.g., Hrdy 1977; Lloyd 1993, 2005). See also the [entry on feminist philosophy of biology](#).

Often but not always, feminist epistemologists go beyond pointing out what they regard as androcentric bias and reject the value-free ideal altogether—with an eye on the social and moral responsibility of scientific inquiry. They try to show that a value-laden science can also meet important criteria for being epistemically reliable and objective (e.g., Anderson 2004; Kourany 2010). A classical representative of such efforts is Longino's (1990) **contextual empiricism**. She reinforces Popper's insistence that "the objectivity of scientific statements lies in the fact that they can be inter-subjectively tested" (1934 [2002]: 22), but unlike Popper, she conceives scientific knowledge essentially as a social product. Thus, our conception of scientific objectivity must directly engage with the social process that generates knowledge. Longino assigns a crucial function to social systems of criticism in securing the epistemic success of science. Specifically, she develops an epistemology which regards a method of inquiry as "objective to the degree that it permits **transformative criticism**" (Longino 1990: 76). For an epistemic community to achieve transformative criticism, there must be:

- **avenues for criticism:** criticism is an essential part of scientific institutions (e.g., peer review);
- **shared standards:** the community must share a set of cognitive values for assessing theories (more on this in [section 3.1](#));
- **uptake of criticism:** criticism must be able to transform scientific practice in the long run;
- **equality of intellectual authority:** intellectual authority must be shared equally among qualified practitioners.

Longino's contextual empiricism can be understood as a development of John Stuart Mill's view that beliefs should never be suppressed, independently of whether they are true or false. Even the most implausible beliefs might be true, and even if they are false, they might contain a grain of truth which is worth preserving or helps to better articulate true beliefs (Mill 1859 [2003: 72]). The underlying intuition is supported by recent empirical research on the epistemic benefits of a diversity of opinions and perspectives (Page 2007). By stressing the social nature of scientific knowledge, and the importance of criticism (e.g., with respect to potential androcentric bias and inclusive practice), Longino's account fits into the broader project of feminist epistemology.

**Standpoint theory** undertakes a more radical attack on traditional scientific objectivity. This view develops Marxist ideas to the effect that epistemic position is related to, and a product of, social position. Feminist standpoint theory builds on these ideas but focuses on gender, racial and other social relations. Feminist standpoint theorists and proponents of "**situated knowledge**" such as Donna Haraway (1988), Sandra Harding (1991, 2015a, 2015b) and Alison Wylie (2003) deny the internal coherence of a view from nowhere: all human knowledge is at base *human* knowledge and therefore necessarily perspectival. But they argue more than that. Not only is perspectivity the human condition, it is also a good thing to have. This is because perspectives, especially the perspectives of underprivileged classes and groups in society, come along with epistemic benefits. These ideas are controversial but they draw attention to the possibility that attempts to rid science of perspectives might not only be futile but also costly: they prevent scientists from having the epistemic benefits certain standpoints afford and from developing knowledge *for* marginalized groups in society. The perspectival stance can also explain why criteria for objectivity often vary with context: the relative importance of epistemic virtues is a matter of goals and interests—in other words, standpoint.

By endorsing a perspectival stance, feminist standpoint theory rejects classical elements of scientific objectivity such as neutrality and impartiality (see [section 3.1](#) above). This is a notable difference to feminist epistemology, which is in principle (though not always in practice) compatible with traditional views of objectivity. Feminist standpoint theory is also a political project. For example, Harding (1991, 1993) demands that scientists, their communities and their practices—in other words, the ways through which knowledge is gained—be investigated as rigorously as the object of knowledge itself. This idea she refers to as “**strong objectivity**” replaces the “weak” conception of objectivity in the empiricist tradition: value-freedom, impartiality, rigorous adherence to methods of testing and inference. Like Feyerabend, Harding integrates a transformation of epistemic standards in science into a broader political project of rendering science more democratic and inclusive. On the other hand, she is exposed to similar objections (see also Haack 2003). Isn't it grossly exaggerated to identify class, race and gender as important factors in the construction of physical theories? Doesn't the feminist approach—like social constructivist approaches—lose sight of the particular epistemic qualities of science? Should non-scientists really have as much authority as trained scientists? To whom does the condition of equally shared intellectual authority apply? Nor is it clear—especially in times of fake news and filter bubbles—whether it is always a good idea to subject scientific results to democratic approval. There is no guarantee (arguably there are few good reasons to believe) that democratized or standpoint-based science leads to more reliable theories, or better decisions for society as a whole.

## 6. Issues in the Special Sciences

So far everything we discussed was meant to apply across all or at least most of the sciences. In this section we will look at a number of specific issues that arise in the social sciences, in economics, and in evidence-based medicine.

### 6.1 Max Weber and Objectivity in the Social Sciences

There is a long tradition in the philosophy of social science maintaining that there is a gulf in terms of both goals as well as methods between the natural and the social sciences. This tradition, associated with thinkers such as the neo-Kantians Heinrich Rickert and Wilhelm Windelband, the hermeneuticist Wilhelm Dilthey, the sociologist-economist Max Weber, and the twentieth-century hermeneuticists Hans-Georg Gadamer and Michael Oakeshott, holds that unlike the natural sciences whose aim it is to establish natural laws and which proceed by experimentation and causal analysis, the social sciences seek understanding (“*Verstehen*”) of social phenomena, the interpretive examination of the meanings individuals attribute to their actions (Weber 1904 [1949]; Weber 1917 [1949]; Dilthey 1910 [1986]; Windelband 1915; Rickert 1929; Oakeshott 1933; Gadamer 1960 [1989]). See also the entries on [hermeneutics](#) and [Max Weber](#).

Understood this way, social science lacks objectivity in more than one sense. One of the more important debates concerning objectivity in the social sciences concerns the role value judgments play and, importantly, whether value-laden research entails claims about the desirability of actions. Max Weber held that the social sciences are necessarily value laden. However, they can achieve some degree of objectivity by keeping out the social researcher's views about whether agents' goals are commendable. In a similar vein, contemporary economics can be said to be value laden because it predicts and explains social phenomena on the basis of agents' preferences. Nevertheless, economists are adamant that economists are not in the business of telling people what they ought to value. Modern economics is thus said to be objective in the Weberian sense of “**absence of researchers' values**”—a conception that we discussed in detail in [section 3](#).

In his widely cited essay “‘Objectivity’ in Social Science and Social Policy” (Weber 1904 [1949]), Weber argued that the idea of an aperspectival social science was meaningless:

There is no absolutely objective scientific analysis of [...] “social phenomena” independent of special and “one-sided” viewpoints according to which expressly or tacitly, consciously or unconsciously they are selected, analyzed and organized for expository purposes. (1904 [1949: 72])

All knowledge of cultural reality, as may be seen, is always knowledge from particular points of view. (1904 [1949: 81])

The reason for this is twofold. First, social reality is too complex to admit of full description and explanation. So we have to select. But, perhaps in contraposition to the natural sciences, we cannot just select those aspects of the phenomena that fall under universal natural laws and treat everything else as “unintegrated residues” (1904 [1949: 73]). This is because, second, in the social sciences we want to understand social phenomena in their individuality, that is, in their unique configurations that have significance for us.

Values solve a selection problem. They tell us what research questions we ought to address because they inform us about the cultural importance of social phenomena:

Only a small portion of existing concrete reality is colored by our value-conditioned interest and it alone is significant to us. It is significant because it reveals relationships which are important to use due to their connection with our values. (1904 [1949: 76])

It is important to note that Weber did not think that social and natural science were different in kind, as Dilthey and others did. Social science too examines the causes of phenomena of interest, and natural science too often seeks to explain natural phenomena in their individual constellations. The role of causal laws is different in the two fields, however. Whereas establishing a causal law is often an end in itself in the natural sciences, in the social sciences laws play an attenuated and accompanying role as mere means to explain cultural phenomena in their uniqueness.

Nevertheless, for Weber social science remains objective in at least two ways. First, once research questions of interest have been settled, answers about the causes of culturally significant phenomena do not depend on the idiosyncrasies of an individual researcher:

But it obviously does not follow from this that research in the cultural sciences can only have results which are “subjective” in the sense that they are valid for one person and not for others. [...] For scientific truth is precisely what is valid for all who seek the truth. (Weber 1904 [1949: 84], emphasis original)

The claims of social science can therefore be objective in our third sense ([see section 4](#)). Moreover, by determining that a given phenomenon is “culturally significant” a researcher reflects on whether or not a practice is “meaningful” or “important”, and not whether or not it is commendable: “Prostitution is a cultural phenomenon just as much as religion or money” (1904 [1949: 81]). An important implication of this view came to the fore in the so-called “*Werturteilsstreit*” (quarrel concerning value judgments) of the early 1900s. In this debate, Weber maintained against the “socialists of the lectern” around Gustav Schmoller the position that social scientists qua scientists should not be directly involved in policy debates because it was not the aim of science to examine the appropriateness of ends. Given a policy goal, a social scientist could make recommendations about effective strategies to reach the goal; but social science was to be value-free in the sense of not taking a stance on the desirability of the goals themselves. This leads us to our conception of objectivity as freedom from value judgments.

## 6.2 Contemporary Rational Choice Theory

Contemporary mainstream economists hold a view concerning objectivity that mirrors Max Weber’s (see above). On the one hand, it is clear that value judgments are at the heart of economic theorizing. “Preferences” are a key concept of rational choice theory, the main theory in contemporary mainstream economics. Preferences are evaluations. If an individual prefers *A* to *B*, she *values A* higher than *B* (Hausman 2012). Thus, to the extent that economists predict and explain market behavior in terms of rational choice theory, they predict and explain market behavior in a way laden with value judgments.

However, economists are not themselves supposed to take a stance about whether or not whatever individuals value is also “objectively” good in a stronger sense:

[...] that an agent is rational from [rational choice theory]’s point of view does not mean that the course of action she will choose is objectively optimal. Desires do not have to align with any objective measure of “goodness”: I may want to risk swimming in a crocodile-infested lake; I may desire to smoke or drink even though I know it harms me. Optimality is determined by the agent’s desires, not the converse. (Paternotte 2011: 307–8)

In a similar vein, Gul and Pesendorfer write:

However, standard economics has no therapeutic ambition, i.e., it does not try to evaluate or improve the individual’s objectives. Economics cannot distinguish between choices that maximize happiness, choices that reflect a sense of duty, or choices that are the response to some impulse. Moreover, standard economics takes no position on the question of which of those objectives the agent should pursue. (Gul and Pesendorfer 2008: 8)

According to the standard view, all that rational choice theory demands is that people’s preferences are (internally) consistent; it has no business in telling people what they ought to prefer, whether their preferences are consistent with external norms or values. Economics is thus value-laden, but laden with the values of the agents whose behavior it seeks to predict and explain and not with the values of those who seek to predict and explain this behavior.

Whether or not social science, and economics in particular, can be objective in this—Weber’s and the contemporary economists’—sense is controversial. On the one hand, there are some reasons to believe that rational choice theory (which is at work not only in economics but also in political science and other social sciences) cannot be applied to empirical phenomena without referring to external norms or values (Sen 1993; Reiss 2013).

On the other hand, it is not clear that economists and other social scientists qua social scientists shouldn’t participate in a debate about social goals. For one thing, trying to do welfare analysis in the standard Weberian way tends to obscure rather than to eliminate normative commitments (Putnam and Walsh 2007). Obscuring value judgments can be detrimental to the social scientist as policy adviser because it will hamper rather than promote trust in social science. For another, economists are in a prime position to contribute to ethical debates, for a variety of reasons, and should therefore take this responsibility seriously (Atkinson 2001).

### **6.3 Evidence-based Medicine and Social Policy**

The same demands calling for “mechanical objectivity” in the natural sciences and quantification in the social and policy sciences in the nineteenth century and mid-twentieth century are responsible for a recent movement in biomedical research, which, even more recently, have swept to contemporary social science and policy. Early proponents of so-called “evidence-based medicine” made their pursuit of a downplay of the “human element” in medicine plain:

Evidence-based medicine de-emphasizes intuition, unsystematic clinical experience, and pathophysiological rationale as sufficient grounds for clinical decision making and stresses the examination of evidence from clinical research. (Guyatt et al. 1992: 2420)

To call the new movement “evidence-based” is a misnomer strictly speaking, as intuition, clinical experience and pathophysiological rationale can certainly constitute evidence. But proponents of evidence-based practices have a much narrower concept of evidence in mind: analyses of the results of randomized controlled trials (RCTs). This movement is now very strong in biomedical research, development economics and a number of areas of social science, especially psychology, education and social policy, and especially in the English speaking world.

The goal is to replace subjective (biased, error-prone, idiosyncratic) judgments by mechanically objective methods. But, as in other areas, attempting to mechanize inquiry can lead to reduced accuracy and utility of the results.

Causal relations in the social and biomedical sciences hold on account of highly complex arrangements of factors and conditions. Whether for instance a substance is toxic depends on details of the metabolic system of the population ingesting it, and whether an educational policy is effective on the constellation of factors that affect the students' learning progress. If an RCT was conducted successfully, the conclusion about the effectiveness of the treatment (or toxicity of a substance) under test is certain for the particular arrangement of factors and conditions of the trial (Cartwright 2007). But unlike the RCT itself, many of whose aspects can be (relatively) mechanically implemented, applying the result to a new setting (recommending a treatment to a patient, for instance) always involves subjective judgments of the kind proponents of evidence-based practices seek to avoid—such as judgments about the similarity of the test to the target or policy population.

On the other hand, RCTs can be regarded as “debiasing procedure” because they prevent researchers from allocating treatments to patients according to their personal interests, so that the healthiest (or smartest or...) subjects get the researcher's favorite therapy. While unbalanced allocations can certainly happen by chance, randomization still provides some warrant that the allocation was not done *on purpose* with a view to promoting somebody's interests. *A priori*, the experimental procedure is thus more impartial with respect to the interests at stake. It has thus been argued that RCTs in medicine, while no guarantor of the best outcomes, were adopted by the U.S. Food and Drugs Administration (FDA) to different degrees during the 1960s and 1970s in order to regain public trust in its decisions about treatments, which it had lost due to the thalidomide and other scandals (Teira and Reiss 2013; Teira 2010). It is important to notice, however, that randomization is at best effective with respect to one kind of bias, viz. selection bias. Important other epistemic concerns are not addressed by the procedure but should not be ignored (Worrall 2002).

## 7. The Unity and Disunity of Scientific Objectivity

In sections 2–5, we have encountered various concepts of scientific objectivity and their limitations. This prompts the question of how unified (or disunified) scientific objectivity is as a concept: Is there something substantive shared by all of these analyses? Or is objectivity, as Heather Douglas (2004) puts it, an “irreducibly complex” concept?

Douglas defends **pluralism about scientific objectivity** and distinguishes three areas of application of the concept: (1) interaction of humans with the world, (2) individual reasoning processes, (3) social processes in science. Within each area, there are various distinct senses which are again irreducible to each other and do not have a common core meaning. This does not mean that the senses are unrelated; they share a complex web of relationships and can also support each other—for example, eliminating values from reasoning may help to achieve procedural objectivity. For Douglas, reducing objectivity to a single core meaning would be a simplification without benefits; instead of a complex web of relations between different senses of objectivity we would obtain an impoverished concept out of touch with scientific practice. Similar arguments and pluralist accounts can be found in Megill (1994), Janack (2002) and Padovani et al. (2015)—see also Axtell (2016).

It has been argued, however, that pluralist approaches give up too quickly on the idea that the different senses of objectivity share one or several important common elements. As we have seen in section [4.1](#) and [5.1](#), scientific objectivity and **trust in science** are closely connected. Scientific objectivity is desirable because to the extent that science is objective we have reasons trust scientists, their results and recommendations (cf. Fine 1998: 18). Thus, perhaps what is unifying among the difference senses of objectivity is that each sense describes a feature of scientific practice that is able to inspire trust in science.

Building on this idea, Inkeri Koskinen has recently argued that it is in fact not trust but reliance that we are after (Koskinen forthcoming). Trust is something that can be betrayed, but only individuals can betray whereas objectivity pertains to institutions, practices, results, etc. We call scientific institutions, practices, results, etc. objective to the extent that we have reasons to rely on them. The analysis does not stop here, however. There is a distinct view about objectivity that is behind Daston and Galison's historical epistemology of the concept and has been defended by Ian Hacking: that objectivity is not a—positive—virtue but rather the absence of this or that vice (Hacking 2015: 26). Speaking of objectivity in imaging, for instance, Daston and Galison write that the goal is to

let the specimen appear without that distortion characteristic of the observer's personal tastes, commitments, or ambitions. (Daston and Galison 2007: 121)

Koskinen picks up this idea of **objectivity as absence of vice** and argues that it is specifically the aversion of *epistemic risks* for which the term is reserved. Epistemic risks comprise “any risk of epistemic error that arises anywhere during knowledge practices” (Biddle and Kukla 2017: 218) such as the risk of having mistaken beliefs, the risk of errors in reasoning and risks related to operationalization, concept formation, and model choice. Koskinen argues that only those epistemic risks that relate to failings of scientists as human beings are relevant to objectivity (Koskinen forthcoming: 13):

For instance, when the results of an experiment are incorrect because of malfunctioning equipment, we do not worry about objectivity—we just say that the results should not be taken into account. [...] So it is only when the epistemic risk is related to our own failings, and is hard to avert, that we start talking about objectivity. Illusions, subjectivity, idiosyncrasies, and collective biases are important epistemic risks arising from our imperfections as epistemic agents.

Koskinen understands her account as a response to Hacking's (2015) criticism that we should stop talking about objectivity altogether. According to Hacking, “objectivity” is an “elevator” or second-level word, similar to “true” or “real”—“Instead of saying that the cat is on the mat, we move up one story and and say that it is true that the cat is on the mat” (2015: 20). He recommends to stick to ground-level questions and worry about whether specific sources of error have been controlled. (A similar elimination request with respect to the labels “objective” and “subjective” in statistical inference has been advanced by Gelman and Hennig (2017).) In focussing on averting specific epistemic risks, Koskinen's account does precisely that. Koskinen argues that a unified account of objectivity as averting epistemic risks takes into account Hacking's negative stance and explains at the same time important features of the concept—for example, why objectivity does not imply certainty and why it varies with context.

The strong point of this account is that none of the threats to a peculiar analysis puts scientific objectivity at risk. We can (and in fact, we do) rely on scientific practices that represent the world from a perspective and where non-epistemic values affect outcomes and decisions. What is left open by Koskinen's account is the normative question of what a scientist who cares about her experiments and inferences being objective should actually do. That is, the philosophical ideas we have reviewed in this section stay mainly on the descriptive level and do not give an actual guideline for working scientists. Connecting the abstract philosophical analysis to day-to-day work in science remains an open problem.

## 8. Conclusions

So is scientific objectivity desirable? Is it attainable? That, as we have seen, depends crucially on how the term is understood. We have looked in detail at four different conceptions of scientific objectivity: faithfulness to facts, value-freedom, freedom from personal biases, and features of community practices. In each case, there are at least some reasons to believe that either science cannot deliver full objectivity in this sense, or that it would not be a good thing to try to do so, or both. Does this mean we should give up the idea of objectivity in science?

We have shown that it is hard to define scientific objectivity in terms of a view from nowhere, value freedom, or freedom from personal bias. It is a lot harder to say anything positive about the matter. Perhaps it is related to a thorough critical attitude concerning claims and findings, as Popper thought. Perhaps it is the fact that many voices are heard, equally respected and subjected to accepted standards, as Longino defends. Perhaps it is something else altogether, or a combination of several factors discussed in this article.

However, one should not (as yet) throw out the baby with the bathwater. Like those who defend a particular explication of scientific objectivity, the critics struggle to explain what makes science objective, trustworthy and special. For instance, our discussion of the value-free ideal (VFI) revealed that alternatives to the VFI are at least as problematic as the VFI itself, and that the VFI may, with all its inadequacies, still be a useful heuristic for fostering scientific integrity and objectivity. Similarly, although entirely “unbiased” scientific procedures

may be impossible, there are many mechanisms scientists can adopt for protecting their reasoning against undesirable forms of bias, e.g., choosing an appropriate method of statistical inference, being transparent about different stages of the research process and avoiding certain questionable research practices.

Whatever it is, it should come as no surprise that finding a positive characterization of what makes science objective is hard. If we knew an answer, we would have done no less than solve the problem of induction (because we would know what procedures or forms of organization are responsible for the success of science). Work on this problem is an ongoing project, and so is the quest for understanding scientific objectivity.

## Bibliography

- Anderson, Elizabeth, 2004, “Uses of Value Judgments in Science: A General Argument, with Lessons from a Case Study of Feminist Research on Divorce”, *Hypatia*, 19(1): 1–24. doi:10.1111/j.1527-2001.2004.tb01266.x
- Atkinson, Anthony B., 2001, “The Strange Disappearance of Welfare Economics”, *Kyklos*, 54(2-3): 193–206. doi:10.1111/1467-6435.00148
- Axtell, Guy, 2016, *Objectivity*, Cambridge: Polity Press.
- Bakker, Marjan, Annette van Dijk, and Jelte M. Wicherts, 2012, “The Rules of the Game Called Psychological Science”, *Perspectives on Psychological Science*, 7(6): 543–554. doi:10.1177/1745691612459060
- Bernardo, J.M., 2012, “Integrated Objective Bayesian Estimation and Hypothesis Testing”, in *Bayesian Statistics 9: Proceedings of the Ninth Valencia Meeting*, J.M. Bernardo et al. (eds.), Oxford: Oxford University Press, 1–68.
- Betz, Gregor, 2013, “In Defence of the Value Free Ideal”, *European Journal for Philosophy of Science*, 3(2): 207–220. doi:10.1007/s13194-012-0062-x
- Biddle, Justin B., 2013, “State of the Field: Transient Underdetermination and Values in Science”, *Studies in History and Philosophy of Science Part A*, 44(1): 124–133. doi:10.1016/j.shpsa.2012.09.003
- Biddle, Justin B. and Rebecca Kukla, 2017, “The Geography of Epistemic Risk”, in *Exploring Inductive Risk: Case Studies of Values in Science*, Kevin C. Elliott and Ted Richards (eds.), New York: Oxford University Press, 215–238.
- Bloor, David, 1982, “Durkheim and Mauss Revisited: Classification and the Sociology of Knowledge”, *Studies in History and Philosophy of Science Part A*, 13(4): 267–297. doi:10.1016/0039-3681(82)90012-7
- Braithwaite, R. B., 1953, *Scientific Explanation*, Cambridge: Cambridge University Press.
- Carnap, Rudolf, 1950 [1962], *Logical Foundations of Probability*, second edition, Chicago: University of Chicago Press.
- Cartwright, Nancy, 2007, “Are RCTs the Gold Standard?”, *BioSocieties*, 2(1): 11–20. doi:10.1017/S1745855207005029
- Chang, Hasok, 2004, *Inventing Temperature: Measurement and Scientific Progress*, Oxford: Oxford University Press. doi:10.1093/0195171276.001.0001
- Churchman, C. West, 1948, *Theory of Experimental Inference*, New York: Macmillan.
- Collins, H. M., 1985, *Changing Order: Replication and Induction in Scientific Practice*, Chicago, IL: University of Chicago Press.
- , 1994, “A Strong Confirmation of the Experimenters’ Regress”, *Studies in History and Philosophy of Science Part A*, 25(3): 493–503. doi:10.1016/0039-3681(94)90063-9
- Cranor, Carl F., 1993, *Regulating Toxic Substances: A Philosophy of Science and the Law*, New York: Oxford University Press.
- Crasnow, Sharon, 2013, “Feminist Philosophy of Science: Values and Objectivity: Feminist Philosophy of Science”, *Philosophy Compass*, 8(4): 413–423. doi:10.1111/phc3.12023
- Daston, Lorraine, 1992, “Objectivity and the Escape from Perspective”, *Social Studies of Science*, 22(4): 597–618. doi:10.1177/030631292022004002
- Daston, Lorraine and Peter Galison, 1992, “The Image of Objectivity”, *Representations*, 40(special issue: Seeing Science): 81–128. doi:10.2307/2928741
- , 2007, *Objectivity*, Cambridge, MA: MIT Press.

- Dilthey, Wilhelm, 1910 [1981], *Der Aufbau der geschichtlichen Welt in den Geisteswissenschaften*, Frankfurt am Main: Suhrkamp.
- Dorato, Mauro, 2004, “Epistemic and Nonepistemic Values in Science”, in Machamer and Wolters 2004: 52–77.
- Douglas, Heather E., 2000, “Inductive Risk and Values in Science”, *Philosophy of Science*, 67(4): 559–579. doi:10.1086/392855
- , 2004, “The Irreducible Complexity of Objectivity”, *Synthese*, 138(3): 453–473. doi:10.1023/B:SYNT.0000016451.18182.91
- , 2009, *Science, Policy, and the Value-Free Ideal*, Pittsburgh, PA: University of Pittsburgh Press.
- , 2011, “Facts, Values, and Objectivity”, Jarvie and Zamora Bonilla 2011: 513–529.
- Duhem, Pierre Maurice Marie, 1906 [1954], *La théorie physique. Son objet et sa structure*, Paris: Chevalier et Riviere; translated by Philip P. Wiener, *The Aim and Structure of Physical Theory*, Princeton, NJ: Princeton University Press, 1954.
- Dupré, John, 2007, “Fact and Value”, in Kincaid, Dupré, and Wylie 2007: 24–71.
- Earman, John, 1992, *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, Cambridge/MA: The MIT Press.
- Elliott, Kevin C., 2011, “Direct and Indirect Roles for Values in Science”, *Philosophy of Science*, 78(2): 303–324. doi:10.1086/659222
- Feyerabend, Paul K., 1962, “Explanation, Reduction and Empiricism”, in H. Feigl and G. Maxwell (ed.), *Scientific Explanation, Space, and Time*, (Minnesota Studies in the Philosophy of Science, 3), Minneapolis, MN: University of Minnesota Press, pp. 28–97.
- , 1975, *Against Method*, London: Verso.
- , 1978, *Science in a Free Society*, London: New Left Books.
- Fine, Arthur, 1998, “The Viewpoint of No-One in Particular”, *Proceedings and Addresses of the American Philosophical Association*, 72(2): 7. doi:10.2307/3130879
- Fisher, Ronald Aylmer, 1935, *The Design of Experiments*, Edinburgh: Oliver and Boyd.
- , 1956, *Statistical Methods and Scientific Inference*, New York: Hafner.
- Franklin, Allan, 1994, “How to Avoid the Experimenters’ Regress”, *Studies in History and Philosophy of Science Part A*, 25(3): 463–491. doi:10.1016/0039-3681(94)90062-0
- , 1997, “Calibration”, *Perspectives on Science*, 5(1): 31–80.
- Freese, Jeremy and David Peterson, 2018, “The Emergence of Statistical Objectivity: Changing Ideas of Epistemic Vice and Virtue in Science”, *Sociological Theory*, 36(3): 289–313. doi:10.1177/0735275118794987
- Gadamer, Hans-Georg, 1960 [1989], *Wahrheit und Methode*, Tübingen : Mohr. Translated as *Truth and Method*, 2<sup>nd</sup> edition, Joel Weinsheimer and Donald G. Marshall (trans), New York, NY: Crossroad, 1989.
- Gelman, Andrew and Christian Hennig, 2017, “Beyond Subjective and Objective in Statistics”, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 180(4): 967–1033. doi:10.1111/rssa.12276
- Giere, Ronald N., 2006, *Scientific Perspectivism*, Chicago, IL: University of Chicago Press.
- Good, Irving John, 1950, *Probability and the Weighing of Evidence*, London: Charles Griffin.
- Gul, Faruk and Wolfgang Pesendorfer, 2008, “The Case for Mindless Economics”, in *The Foundations of Positive and Normative Economics: a Handbook*, Andrew Caplin and Andrew Schotter (eds), New York, NY: Oxford University Press, pp. 3–39.
- Guyatt, Gordon, John Cairns, David Churchill, Deborah Cook, Brian Haynes, Jack Hirsh, Jan Irvine, Mark Levine, Mitchell Levine, Jim Nishikawa, et al., 1992, “Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine”, *JAMA: The Journal of the American Medical Association*, 268(17): 2420–2425. doi:10.1001/jama.1992.03490170092032
- Haack, Susan, 2003, *Defending Science—Within Reason: Between Scientism and Cynicism*, Amherst, NY: Prometheus Books.
- Hacking, Ian, 1965, *Logic of Statistical Inference*, Cambridge: Cambridge University Press. doi:10.1017/CBO9781316534960
- , 2015, “Let’s Not Talk About Objectivity”, in Padovani, Richardson, and Tsou 2015: 19–33. doi:10.1007/978-3-319-14349-1\_2
- Hanson, Norwood Russell, 1958, *Patterns of Discovery: An Inquiry into the Conceptual Foundations of Science*, Cambridge: Cambridge University Press.

- Haraway, Donna, 1988, "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective", *Feminist Studies*, 14(3): 575–599. doi:10.2307/3178066
- Harding, Sandra, 1991, *Whose Science? Whose Knowledge? Thinking from Women's Lives*, Ithaca, NY: Cornell University Press.
- , 1993, "Rethinking Standpoint Epistemology: What is Strong Objectivity?", in *Feminist Epistemologies*, Linda Alcoff and Elizabeth Potter (ed.), New York, NY: Routledge, 49–82.
- , 2015a, *Objectivity and Diversity: Another Logic of Scientific Research*, Chicago: University of Chicago Press.
- , 2015b, "After Mr. Nowhere: What Kind of Proper Self for a Scientist?", *Feminist Philosophy Quarterly*, 1(1): 1–22. doi:10.5206/fpq/2015.1.2
- Hausman, Daniel M., 2012, *Preference, Value, Choice, and Welfare*, New York: Cambridge University Press. doi:10.1017/CBO9781139058537
- Hempel, Carl G., 1965, *Aspects of Scientific Explanation*, New York: The Free Press.
- Hesse, Mary B., 1980, *Revolutions and Reconstructions in the Philosophy of Science*, Bloomington, IN: University of Indiana Press.
- Howson, Colin, 2000, *Hume's Problem: Induction and the Justification of Belief*, Oxford: Oxford University Press.
- Howson, Colin and Peter Urbach, 1993, *Scientific Reasoning: The Bayesian Approach*, second edition, La Salle, IL: Open Court.
- Hrdy, Sarah Blaffer, 1977, *The Langurs of Abu: Female and Male Strategies of Reproduction*, Cambridge, MA: Harvard University Press.
- Ioannidis, John P. A., 2005, "Why Most Published Research Findings Are False", *PLoS Medicine*, 2(8): e124. doi:10.1371/journal.pmed.0020124
- Janack, Marianne, 2002, "Dilemmas of Objectivity", *Social Epistemology*, 16(3): 267–281. doi:10.1080/0269172022000025624
- Jarvie, Ian C. and Jesús P. Zamora Bonilla (eds.), 2011, *The SAGE Handbook of the Philosophy of Social Sciences*, London: SAGE.
- Jaynes, Edwin T., 1968, "Prior Probabilities", *IEEE Transactions on Systems Science and Cybernetics*, 4(3): 227–241. doi:10.1109/TSSC.1968.300117
- Jeffrey, Richard C., 1956, "Valuation and Acceptance of Scientific Hypotheses", *Philosophy of Science*, 23(3): 237–246. doi:10.1086/287489
- Jeffreys, Harold, 1939 [1980], *Theory of Probability*, third edition, Oxford: Oxford University Press.
- Kelvin, Lord (William Thomson), 1883, "Electrical Units of Measurement", Lecture to the Institution of Civil Engineers on 3 May 1883, reprinted in 1889, *Popular Lectures and Addresses*, Vol. I, London: MacMillan and Co., p. 73.
- Kincaid, Harold, John Dupré, and Alison Wylie (eds.), 2007, *Value-Free Science?: Ideals and Illusions*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780195308969.001.0001
- Kitcher, Philip, 2011a, *Science in a Democratic Society*, Amherst, NY: Prometheus Books.
- , 2011b, *The Ethical Project*, Cambridge, MA: Harvard University Press.
- Koskinen, Inkeri, forthcoming, "Defending a Risk Account of Scientific Objectivity", *The British Journal for the Philosophy of Science*, first online: 3 August 2018. doi:10.1093/bjps/axy053
- Kourany, Janet A., 2010, *Philosophy of Science after Feminism*, Oxford: Oxford University Press.
- Kuhn, Thomas S., 1962 [1970], *The Structure of Scientific Revolutions*, second edition, Chicago: University of Chicago Press.
- , 1977, "Objectivity, Value Judgment, and Theory Choice", in his *The Essential Tension. Selected Studies in Scientific Tradition and Change*, Chicago: University of Chicago Press: 320–39.
- Lacey, Hugh, 1999, *Is Science Value-Free? Values and Scientific Understanding*, London: Routledge.
- , 2002, "The Ways in Which the Sciences Are and Are Not Value Free", in *In the Scope of Logic, Methodology and Philosophy of Science: Volume Two of the 11th International Congress of Logic, Methodology and Philosophy of Science, Cracow, August 1999*, Peter Gärdenfors, Jan Woleński, and Katarzyna Kijania-Placek (eds.), Dordrecht: Springer Netherlands, 519–532. doi:10.1007/978-94-017-0475-5\_9
- Laudan, Larry, 1984, *Science and Values: An Essay on the Aims of Science and Their Role in Scientific Debate*, Berkeley/Los Angeles: University of California Press.

- Levi, Isaac, 1960, "Must the Scientist Make Value Judgments?", *The Journal of Philosophy*, 57(11): 345–357. doi:10.2307/2023504
- Lloyd, Elisabeth A., 1993, "Pre-Theoretical Assumptions in Evolutionary Explanations of Female Sexuality", *Philosophical Studies*, 69(2–3): 139–153. doi:10.1007/BF00990080
- , 2005, *The Case of the Female Orgasm: Bias in the Science of Evolution*, Cambridge, MA: Harvard University Press.
- Longino, Helen E., 1990, *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*, Princeton, NY: Princeton University Press.
- , 1996, "Cognitive and Non-Cognitive Values in Science: Rethinking the Dichotomy", in *Feminism, Science, and the Philosophy of Science*, Lynn Hankinson Nelson and Jack Nelson (eds.), Dordrecht: Springer Netherlands, 39–58. doi:10.1007/978-94-009-1742-2\_3
- Machamer, Peter and Gereon Wolters (eds.), 2004, *Science, Values and Objectivity*, Pittsburgh: Pittsburgh University Press.
- Mayo, Deborah G., 1996, *Error and the Growth of Experimental Knowledge*, Chicago & London: The University of Chicago Press.
- McMullin, Ernan, 1982, "Values in Science", *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1982*, 3–28.
- , 2009, "The Virtues of a Good Theory", in *The Routledge Companion to Philosophy of Science*, Martin Curd and Stathis Psillos (eds), London: Routledge.
- Megill, Allan, 1994, "Introduction: Four Senses of Objectivity", in *Rethinking Objectivity*, Allan Megill (ed.), Durham, NC: Duke University Press, 1–20.
- Mill, John Stuart, 1859 [2003], *On Liberty*, New Haven and London: Yale University Press.
- Mitchell, Sandra D., 2004, "The Prescribed and Proscribed Values in Science Policy", in Machamer and Wolters 2004: 245–255.
- Nagel, Thomas, 1986, *The View From Nowhere*, New York, NY: Oxford University Press.
- Nixon, Richard, 1969, "Special Message to the Congress on Social Security", 25 September 1969. [[Nixon 1969 available online](#)]
- Norton, John D., 2003, "A Material Theory of Induction", *Philosophy of Science*, 70(4): 647–670. doi:10.1086/378858
- , 2008, "Must Evidence Underdetermine Theory?", in *The Challenge of the Social and the Pressure of Practice*, Martin Carrier, Don Howard and Janet Kourany (eds), Pittsburgh, PA: Pittsburgh University Press: 17–44.
- Oakeshott, Michael, 1933, *Experience and Its Modes*, Cambridge: Cambridge University Press.
- Okruhlik, Kathleen, 1994, "Gender and the Biological Sciences", *Canadian Journal of Philosophy Supplementary Volume*, 20: 21–42. doi:10.1080/00455091.1994.10717393
- Open Science Collaboration, 2015, "Estimating the Reproducibility of Psychological Science", *Science*, 349(6251): aac4716. doi:10.1126/science.aac4716
- Padovani, Flavia, Alan Richardson, and Jonathan Y. Tsou (eds.), 2015, *Objectivity in Science: New Perspectives from Science and Technology Studies*, (Boston Studies in the Philosophy and History of Science 310), Cham: Springer International Publishing. doi:10.1007/978-3-319-14349-1
- Page, Scott E., 2007, *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*, Princeton, NJ: Princeton University Press.
- Paternotte, Cédric, 2011, "Rational Choice Theory", in *The SAGE Handbook of The Philosophy of Social Sciences*, Jarvie and Zamora Bonilla 2011: 307–321.
- Popper, Karl. R., 1934 [2002], *Logik der Forschung*, Vienna: Julius Springer. Translated as *Logic of Scientific Discovery*, London: Routledge.
- , 1963, *Conjectures and Refutations: The Growth of Scientific Knowledge*, New York: Harper.
- , 1972, *Objective Knowledge: An Evolutionary Approach*, Oxford: Oxford University Press.
- Porter, Theodore M., 1995, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*, Princeton, NJ, Princeton University Press.
- Putnam, Hilary, 2002, *The Collapse of the Fact/Value Dichotomy and Other Essays*, Cambridge, MA: Harvard University Press.
- Putnam, Hilary and Vivian Walsh, 2007, "A Response to Dasgupta", *Economics and Philosophy*, 23(3): 359–364. doi:10.1017/S026626710700154X

- Reichenbach, Hans, 1938, "On Probability and Induction", *Philosophy of Science*, 5(1): 21–45. doi:10.1086/286483
- Reiss, Julian, 2008, *Error in Economics: The Methodology of Evidence-Based Economics*, London: Routledge.
- , 2010, "In Favour of a Millian Proposal to Reform Biomedical Research", *Synthese*, 177(3): 427–447. doi:10.1007/s11229-010-9790-7
- , 2013, *Philosophy of Economics: A Contemporary Introduction*, New York, NY: Routledge.
- , 2020, "What Are the Drivers of Induction? Towards a Material Theory+", *Studies in History and Philosophy of Science Part A* 83: 8–16.
- Resnik, David B., 2007, *The Price of Truth: How Money Affects the Norms of Science*, Oxford: Oxford University Press.
- Rickert, Heinrich, 1929, *Die Grenzen der naturwissenschaftlichen Begriffsbildung. Eine logische Einleitung in die historischen Wissenschaften*, 6th edition, Tübingen: Mohr Siebeck. First edition published in 1902.
- Royall, Richard, 1997, *Scientific Evidence: A Likelihood Paradigm*, London: Chapman & Hall.
- Rudner, Richard, 1953, "The Scientist *qua* Scientist Makes Value Judgments", *Philosophy of Science*, 20(1): 1–6. doi:10.1086/287231
- Ruphy, Stéphanie, 2006, "'Empiricism All the Way down': A Defense of the Value-Neutrality of Science in Response to Helen Longino's Contextual Empiricism", *Perspectives on Science*, 14(2): 189–214. doi:10.1162/posc.2006.14.2.189
- Sen, Amartya, 1993, "Internal Consistency of Choice", *Econometrica*, 61(3): 495–521.
- Shrader-Frechette, K. S., 1991, *Risk and Rationality*, Berkeley/Los Angeles: University of California Press.
- Simonsohn, Uri, Leif D. Nelson, and Joseph P. Simmons, 2014, "P-Curve: A Key to the File-Drawer.", *Journal of Experimental Psychology: General*, 143(2): 534–547. doi:10.1037/a0033242
- Sprenger, Jan, 2016, "Bayesianism vs. Frequentism in Statistical Inference", in *Oxford Handbook on Philosophy of Probability*, Alan Hájek and Christopher Hitchcock (eds), Oxford: Oxford University Press.
- , 2018, "The Objectivity of Subjective Bayesianism", *European Journal for Philosophy of Science*, 8(3): 539–558. doi:10.1007/s13194-018-0200-1
- Sprenger, Jan and Stephan Hartmann, 2019, *Bayesian Philosophy of Science*, Oxford: Oxford University Press. doi:10.1093/oso/9780199672110.001.0001
- Steel, Daniel, 2010, "Epistemic Values and the Argument from Inductive Risk", *Philosophy of Science*, 77(1): 14–34. doi:10.1086/650206
- Steele, Katie, 2012, "The Scientist *qua* Policy Advisor Makes Value Judgments", *Philosophy of Science*, 79(5): 893–904. doi:10.1086/667842
- Stegenga, Jacob, 2011, "Is Meta-Analysis the Platinum Standard of Evidence?", *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 42(4): 497–507. doi:10.1016/j.shpsc.2011.07.003
- , 2018, *Medical Nihilism*, Oxford: Oxford University Press. doi:10.1093/oso/9780198747048.001.0001
- Teira, David, 2010, "Frequentist versus Bayesian Clinical Trials", in *Philosophy of Medicine*, Fred Gifford (ed.), (Handbook of the Philosophy of Science 16), Amsterdam: Elsevier, 255–297. doi:10.1016/B978-0-444-51787-6.50010-6
- Teira, David and Julian Reiss, 2013, "Causality, Impartiality and Evidence-Based Policy", in *Mechanism and Causality in Biology and Economics*, Hsiang-Ke Chao, Szu-Ting Chen, and Roberta L. Millstein (eds.), (History, Philosophy and Theory of the Life Sciences 3), Dordrecht: Springer Netherlands, 207–224. doi:10.1007/978-94-007-2454-9\_11
- Weber, Max, 1904 [1949], "Die 'Objektivität' sozialwissenschaftlicher und sozialpolitischer Erkenntnis", *Archiv für Sozialwissenschaft und Sozialpolitik*, 19(1): 22–87. Translated as "'Objectivity' in Social Science and Social Policy", in Weber 1949: 50–112.
- , 1917 [1949], "Der Sinn der 'Wertfreiheit' der soziologischen und ökonomischen Wissenschaften". Reprinted in *Gesammelte Aufsätze zur Wissenschaftslehre*, Tübingen: UTB, 1988, 451–502. Translated as "The Meaning of 'Ethical Neutrality' in Sociology and Economics" in Weber 1949: 1–49.
- , 1949, *The Methodology of the Social Sciences*, Edward A. Shils and Henry A. Finch (trans/eds), New York, NY: Free Press.
- Wilholt, Torsten, 2009, "Bias and Values in Scientific Research", *Studies in History and Philosophy of Science Part A*, 40(1): 92–101. doi:10.1016/j.shpsa.2008.12.005


- , 2013, “Epistemic Trust in Science”, *The British Journal for the Philosophy of Science*, 64(2): 233–253. doi:10.1093/bjps/axs007
- Williams, Bernard, 1985 [2011], *Ethics and the Limits of Philosophy*, Cambridge, MA: Harvard University Press. Reprinted London and New York, NY: Routledge, 2011.
- Williamson, Jon, 2010, *In Defence of Objective Bayesianism*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199228003.001.0001
- Windelband, Wilhelm, 1915, *Präludien. Aufsätze und Reden zur Philosophie und ihrer Geschichte*, fifth edition, Tübingen: Mohr Siebeck.
- Winsberg, Eric, 2012, “Values and Uncertainties in the Predictions of Global Climate Models”, *Kennedy Institute of Ethics Journal*, 22(2): 111–137. doi:10.1353/ken.2012.0008
- Wittgenstein, Ludwig, 1953 [2001], *Philosophical Investigations*, G. Anscombe (trans.), London: Blackwell.
- Worrall, John, 2002, “What Evidence in Evidence-Based Medicine?”, *Philosophy of Science*, 69(S3): S316–S330. doi:10.1086/341855
- Wylie, Alison, 2003, “Why Standpoint Matters”, in *Science and Other Cultures: Issues in Philosophies of Science and Technology*, Robert Figueroa and Sandra Harding (eds), New York, NY and London: Routledge, pp. 26–48.
- Ziliak, Stephen Thomas and Deirdre N. McCloskey, 2008, *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice and Lives*, Ann Arbor, MI: University of Michigan Press.

## Academic Tools

 [How to cite this entry.](#)

 [Preview the PDF version of this entry](#) at the [Friends of the SEP Society](#).

 [Look up this entry topic](#) at the [Internet Philosophy Ontology Project](#) (InPhO).

 [Enhanced bibliography for this entry](#) at [PhilPapers](#), with links to its database.

## Other Internet Resources

- Norton, John, manuscript, [The Material Theory of Induction](#), retrieved on 9 January 2020.
- [Objectivity](#), entry by Dwayne H. Mulder in the *Internet Encyclopedia of Philosophy*.

## Related Entries

[Bayes' Theorem](#) | [confirmation](#) | [feminist philosophy, interventions: epistemology and philosophy of science](#) | [feminist philosophy, interventions: philosophy of biology](#) | [Feyerabend, Paul](#) | [hermeneutics](#) | [incommensurability: of scientific theories](#) | [Kuhn, Thomas](#) | [logic: inductive](#) | [physics: experiment in](#) | [science: theory and observation in](#) | [scientific realism](#) | [statistics, philosophy of](#) | [underdetermination, of scientific theories](#) | [Weber, Max](#)

Copyright © 2020 by  
[Julian Reiss](#) <[julian.reiss@jku.at](mailto:julian.reiss@jku.at)>  
[Jan Sprenger](#) <[jan.sprenger@unito.it](mailto:jan.sprenger@unito.it)>

[Open access to the SEP is made possible by a world-wide funding initiative.](#)  
[Please Read How You Can Help Keep the Encyclopedia Free](#)

**Stanford** | Center for the Study of  
 Language and Information

The Stanford Encyclopedia of Philosophy is [copyright © 2020](#) by [The Metaphysics Research Lab](#), Center for the Study of Language and Information (CSLI), Stanford University

Library of Congress Catalog Data: ISSN 1095-5054