

Catching the WAVE

The Weight-Adjusting Account of Values and Evidence

Boaz Miller

Forthcoming in *Studies in History and Philosophy of Science*

Abstract It is commonly argued that values “fill the logical gap” of underdetermination of theory by evidence, namely, values affect our choice between two or more theories that fit the same evidence. The underdetermination model, however, does not exhaust the roles values play in evidential reasoning. I introduce WAVE – a novel account of the logical relations between values and evidence. WAVE states that values influence evidential reasoning by adjusting evidential weights. I argue that the weight-adjusting role of values is distinct from their underdetermination gap-filling role. Values adjust weights in three ways. First, values affect our trust in the testimony of others. Second, values influence the evidential thresholds required for justified epistemic judgments. Third, values influence the relative weight of a certain type of evidence within a body of multimodal discordant evidence. WAVE explains, from an epistemic perspective, rather than psychological, how smokers, for example, can find the same evidence about the dangers of smoking less persuasive than non-smokers. WAVE allows for a wider effect of values on our accepted scientific theories and beliefs than the effect for which the underdetermination model allows alone; therefore, science studies scholars must consider WAVE in their research and analysis of evidential case studies.

Keywords Science and Values · Evidence · Testimony · Trust · Underdetermination · Science and Technology Studies (STS)

There is a renewed interest in the interactions between values and evidence in science. Social epistemologists, philosophers and sociologists of science have examined various ways in which values, particularly social, are involved in evidential reasoning and affect its outcomes. It is unclear, however, how the different roles of values fit within a unified account of the logical relations between values, evidence, and theory or belief. Specifically, how do these roles relate to the thesis that values fill the logical gap of underdetermination of theory by evidence? Do they merely amount to different ways values fill the underdetermination gap, or should they be characterized differently?

I introduce the Weight-Adjusting Account of Values and Evidence (WAVE). WAVE states that values influence evidential reasoning by adjusting evidential weights. The weight-adjusting role of values is distinct from their underdetermination gap-filling role. WAVE supplements the underdetermination model, which only partly captures the logical relations between values, evidence, and theory or belief.

That some roles values play in evidential reasoning are not manifestations of the gap-filling role has been alluded to in the Science and Values literature. However, it has not been explicitly argued for, which has been cause for confusion. For example, an edited volume of top-of-the-art papers about underdetermination from the seventies (Harding, 1976) contains no discussion of, or reference to, the argument from inductive risk. More recently, Douglas (2009), who gives a novel defence of this argument, avoids the language of underdetermination altogether. Some scholars, such as Elliot (2011, pp. 62-70), regard the argument from inductive risk as distinct from underdetermination, while, others, such as Biddle (2013), assume that the underdetermination model encompasses all epistemically relevant roles of values. This paper aims at clearing this confusion by distinguishing the underdetermination gap-filling role from the weight-adjusting role.

WAVE identifies three ways values adjust evidential weights. *First*, values affect the trust people assign to others' testimonies. *Second*, values affect the threshold level evidence

must meet for a justified epistemic judgment. *Third*, values affect the relative weights of different types of evidence within a body of multimodal discordant evidence. WAVE is a descriptive account. WAVE does not distinguish legitimate from illegitimate roles for values, but a normative account can benefit from its distinctions.

Section 1 reviews the underdetermination model and its limitations. Section 2 argues that social values affect the trust we extend to others' testimonies, and that this is not another aspect of filling the underdetermination gap. Section 3 reviews research in experimental psychology on the influence of values on evidential reasoning, which prepares the ground for the next sections. Section 4 argues that social values lower and raise evidential thresholds. Section 5 argues that they affect the relative weighing of discordant evidence. Section 6 applies WAVE to two case studies to illustrate its usefulness to empirical science studies.

1. The Underdetermination Gap-Filling Model and Its Discontents

1.1. The Underdetermination Gap-Filling Model

According to the thesis of underdetermination of theory by evidence, any body of evidence can be logically accommodated by more than one theory and perhaps infinitely many. It is argued that values "fill the gap" between theory and evidence; namely, because evidence alone does not determine theory or fix belief, agents must implicitly or explicitly appeal to values to choose which theory to accept or belief to form. They adopt the theories or beliefs most consonant with the values they cherish. In this section, I briefly present the underdetermination gap-filling model to contrast it with WAVE.

I use a broad notion of values. A value is anything that serves as a basis for discriminating between different states of affairs and ranking some of them higher than others with respect to how much they are desired or cared about or how the personal, social, natural, or cosmic order ought to be (cf. Taylor 1992, pp. 29-30). Political, ideological, and ethical values naturally fall under this rubric, but also social interests and psychological motivations, as they discriminate between more and less desirable states of affairs in our personal and social lives. Theoretical or cognitive values, such as simplicity, scope, and predictive power, fall under this rubric as well, as they rank theories that have certain properties as higher than others in these respects. Similarly, I use a broad notion of evidence as anything that legitimately serves as a basis for discriminating between different descriptions of the world or parts of it, and ranking some of them more probable than others.

Several distinctions are drawn regarding underdetermination. One is between transient and in-principle underdetermination. Transient underdetermination regards a theory's being underdetermined by evidence as relative to a specific time. Its proponents argue that there is no guarantee that an underdetermined theory at a specific time will continue to be so as time goes by and more evidence is gathered (Laudan & Leplin, 1991). By contrast, proponents of in-principle underdetermination, associated with Quine (1951) and Duhem (1954), argue that in principle, evidence alone cannot fix a single theory regardless of the state of knowledge at a given time (Potter, 1996). Transient underdetermination assigns a more restricted role to values in theory choice, which may be diminished or eliminated at a future time. Yet, there can be non-trivial conceptual differences between transiently underdetermined rival theories (Carrier, 2011).

Another distinction is between concrete and potential underdetermination. In concrete underdetermination, the evidence cannot fix a theory out of a set of concrete alternatives, while in potential underdetermination, it cannot fix a theory out of a set that consists *inter alia* of unconsidered theories. Potential underdetermination fuels sceptical arguments about scientific knowledge and objectivity, which state that if scientists choose a theory from a pool of all false

or biased theories, and there are possible preferable yet unconsidered alternatives, then the truth or objectivity of scientific knowledge is questionable (Stanford, 2010; Okruhlik, 1998).

In their gap-filling role, values may influence theory choice *bottom-up* or *top-down*. Bottom-up, values implicitly affect scientists' background assumptions on which they base their theories. Such background assumptions are "neither self-evident nor logically true" (Longino, 2002, p. 128). They guide the interpretation of data *qua* evidence. For example, chipped stones from early human evolution can serve as evidence for male hunting tools, or for female vegetable-processing tools, thus supporting two different theories about the driving forces of human evolution (Longino, 1990, pp. 104-111). Prevailing social values, e.g., sexist values, may implicitly serve as reasons for adopting certain theoretical background assumptions rather than others, and blind researchers to other possible background assumptions that would support another theory. Bottom-up influence thus accords well with potential underdetermination, because scientists' background assumptions restrict the possible theories they consider.

Top-down, values act as "tiebreakers", namely, given two or more empirically equivalent theories, values determine which one is adopted. For example, *ceteris paribus*, if scientists adhere to simplicity, they choose the simpler theory. Top-down influence accords well with concrete underdetermination, because values help choose between actually available alternatives.¹

Top-down influence on theory choice is discussed by Kuhn (1977), who recognizes five epistemic values constitutive of science: accuracy, consistency, scope, simplicity, and fruitfulness.² While Kuhn's (1977) account is sometimes called "Kuhnian underdetermination" (Carrier 2008), I should stress that it does not, strictly speaking, fully conform to the gap-filling model for two reasons. *First*, under the gap-filling model, scientists use values to decide between rival theories that accommodate the same evidence. By contrast, according to Kuhn (1970, p. 85), during a crisis, there is only *partial overlap* between the problems that can be solved by the old and the new paradigms. The new theory can account for evidence the old theory cannot and vice versa.³ *Second*, according to Kuhn, scientists do not always choose a new theory because it can better accommodate the evidence, but because they have *faith in its potential, still unrealized, to eventually do so*:

¹ Intemann (2005, pp. 1008-1010) identifies a third role of values in evidential reasoning and theory choice, which is that values play a constitutive normative role in the construction of theories, i.e., the concepts that are used in the construction of the theories are irreducibly normative (cf. Putnam, 2002). Intemann argues that this role is distinct from filling the underdetermination gap. I disagree with her analysis, and think that this role can be accommodated within the gap-filling model. In my view, this is just a combination of the gap-filling role and the theory-ladenness of observation, i.e., the inability to describe data with concepts that do not presuppose any theoretical commitments. In this case, the theoretical concepts happen to be normative. The problem of the theory ladenness of observation is not unique to normative concepts. Much of the Science and Values literature – this paper included – tends to assume for the sake of simplicity that data and theory are logically distinguishable, whereas in reality, this is rarely, if ever, the case.

² This list may be debated. For example, Longino (1995) lists alternative epistemic values, which reflect a feminist agenda.

³ Cf. Kitcher (2012, p. 207; emphasis in the origin): "Philosophers have been beguiled by the thought that there is a single problem of underdetermination of theory by evidence that affects all areas of science equally. This is doubly mistaken, for there is an important difference between *Quinean* underdetermination (roughly, cases in which rival theories are supposed to receive exactly the same support from the same body of evidence) and *Kuhnian* underdetermination (roughly, cases in which rival theories are successful in rather different ways)".

The man who embraces a new paradigm at an early stage must do so *in defiance of the evidence* provided by problem-solving. He must, that is, have faith that the new paradigm will succeed with the many large problems that confront it [...] A decision of that kind can only be made on faith (1970, pp. 157-158; emphasis added).

According to Kuhn (1970, pp. 66-75), in a crisis, scientists re-rank the evidence that they have with respect to its importance. Minor problems, which were once thought to be capable of being eventually explained, suddenly become pressing counterevidence against the old theory. I will argue in Section 5 that only WAVE adequately characterizes the role of values in this re-ranking process.

Let me distinguish *underdetermination* from mere *undetermination* (or non-determination). *Undetermination* is a weaker relation than *underdetermination*. When *X* is *undetermined* by *Y*, we need not presuppose logical relations between *X* and *Y*. For example, the number of passengers in the bus is *undetermined* (and *undeterminable*) by the age of the driver or the colour of my shirt. By contrast, when a theory *T* is *underdetermined* by evidence *E*, we already assume there are *T* and *E* which stand in semantic and epistemic relations, namely, relations of representation and confirmation. We also assume at least one more possible or concrete theory *T'*, which also stands in such relations with *E*. When values fill the logical gap of underdetermination of theory by evidence, this means that they fix the choice between *T*, *T'* and possibly more theories that presumably stand in similar relations to *E*.

Thus, when I argue that WAVE is distinct from the gap-filling model, I mean that in WAVE, values do not fix theory between two or more possible or concrete theories that stand in the above-specified semantic and epistemic relations with a given body of evidence. Rather, they play a different role, which may even be temporally and logically prior to underdetermination-gap filling. In WAVE, there is no contrastive theory choice, i.e., values *inter alia* determine whether the evidence sufficiently supports a theory or a judgment, regardless of there being other concrete or possible theories that fit the evidence.

1.2. Criticism of the Gap-Filling Model

So far I reviewed the underdetermination thesis and the two ways values fill the logical gap between theory and evidence: top-down, and bottom-up. I clarified the presuppositions of the gap-filling model, and what it means for values to participate in evidential reasoning in ways other than filling this gap. I will now briefly review the debate about the role of values in filling the underdetermination gap, which concerns the *kinds* of values that legitimately fill the gap, and the *extent* values can, do, and should fill it.

Regarding the kinds of values that legitimately fill the gap, some argue that only cognitive values, e.g., simplicity, scope, or explanatory power, may legitimately influence theory choice. Such values are considered benign and internal to science (McMullin, 1983; Laudan, 2004). Others argue that social, political, and ideological values may legitimately fill the gap as well. For example, Kourany (2010, pp. 69-75) argues that just like simplicity or scope, racial equality may constitute a legitimate reason for preferring a theory consonant with it rather than one that is not. This view is contested. Critics argue that the fact that social values play a role in theory choice does not entail that they should. According to this objection, social values reflect our desired social order – how the social world should be, while theories describe how the world is. Hence, social values are external to the aims of science and constitute biases that should be eliminated (Intemann, 2005).

Proponents of a legitimate gap-filling role for social values respond that the above objection presupposes an untenable, sharp, principled, and meaningful distinction between cognitive and social values. But social and cognitive values cannot be sharply distinguished, and

since the critics acknowledge that the influence of *some* values, namely, cognitive values, on theory choice is necessary and benign, then social values may also play a legitimate role in theory choice (Longino, 2002, pp. 77-96, Machamer & Douglas, 1999; Solomon, 2001, pp. 51-63). This does *not* mean, however, that the effect of any social value is legitimate in any context.

Another criticism of the gap argument concerns the *extent* social values fill the gap. Norton writes that the gap argument rests on an “improvised and oversimplified account of the nature of inductive inference” (2008, p. 19). He argues that under most commonly used confirmation models, when two theories accommodate the same evidence, they usually do not enjoy the same inductive support, i.e., the same evidence does not equally confirm them. Usually, the evidence inductively favours one theory over another. Norton therefore argues that social values fill the underdetermination gap only in rare cases in which theories enjoy similar inductive support. Haack (1998, pp. 110-111) adds that even when two theories enjoy the same warrant, scientists need not appeal to social values to choose between them – they can simply withhold judgment.

Critics of the gap argument commonly assume that gap-filling exhausts all the epistemically relevant roles of values. Because they think the influence of values on theory choice for which the gap-filling model allows is restricted, they regard arguments that state that scientific knowledge is socially constructed as hyperbolic and unsubstantiated. Pinnick, for example, states that if feminist epistemologists who make such claims want to get “more than a yawn” (1994, p. 651) from philosophers of science, they should appeal to arguments other than underdetermination. By showing that the underdetermination gap-filling model is far from exhaustive, WAVE should wake up the yawning philosophers from their dogmatic sleep. That is, while critics may point out valid deficiencies of the gap argument, their criticism does not entail that the role of values in evidential justification is limited. In the next section, I start spelling out the weight-adjusting role by arguing that values affect people’s trust in testimony.

2. Social Values Affect Trust in Testimony

Rolin (2004) claims that by focusing on the gap-filling role of values, social epistemologists have neglected values’ effect on trust in testimony. She notes that social values may skew scientists’ assessment of their colleagues’ trustworthiness. For example, sexist values make some male scientists unjustifiably underrate female scientists’ testimonies and overrate males’ testimonies. Rolin does not conclusively show, however, that this influence cannot be accommodated within the gap-filling model. In this section, I argue that it cannot. I argue because testimonial-belief formation does not necessarily involve theory choice, the influence of values on it cannot be another aspect of filling the underdetermination gap.

According to Baier (1986), trust is one’s reliance on another’s good will toward one. Trust differs from mere reliance on another’s behaviour in that it is directed at a person’s good will. When one depends on another’s good will, one becomes vulnerable to its limits. Trust involves one’s accepted vulnerability to another’s possible but unexpected ill will. Baier’s account requires some fine tuning. First, sometimes, only lack of ill will is enough for trust. Second, one extends or denies trust not as reliance merely on another’s good will, but also on his competence – I may not trust Jack with some fragile equipment, not because I suspect his good will, but because I suspect his ability to handle it with care. Indeed, testimonial trust is typically cashed out in terms of reliance on one’s *sincerity* and *competence* (e.g., Fricker, 1995). When one is sincere with me, one shows *epistemic* good will toward me. When I trust one’s testimony, I rely on one’s good will insofar as it relates to my achieving knowledge.⁴

⁴ Epistemic aims may be overwritten by other aims. For example, the person I trust may think that it would be best for my psychological wellbeing that I did not know that my wife was cheating on me, and lie to me about

Returning to the claim that social values affect trust in testimony, one may argue that this is just another way values fill the underdetermination gap, because when a hearer decides whether to trust a speaker's testimony, she in effect constructs and entertains in her mind theories about the speaker's trustworthiness that fit her available evidence. She appeals to values to construct the theories and choose among them. She extends or denies trust based on her chosen theory. Call this "*the theoretical account of trust*". I argue against this account. I need not show that a hearer *never* constructs a theory about the speaker's trustworthiness, but only that she *sometimes* need not.

Two claims underpin the theoretical account of trust: (1) testimonial beliefs are inferential; (2) trust is a propositional attitude. Namely, trust is grounded in a belief that a person is trustworthy, which the subject has inferred from her available evidence. I will argue that arguments for these two claims are inconclusive. Inasmuch as testimonial beliefs are not inferential and trust is not a propositional attitude, the influence of values on trust in testimony is distinct from filling the underdetermination gap.

Thagard (2006) and Lipton (2007) argue that testimonial beliefs are inferential. They suggest models in which a person's trust in a speaker's testimony is based on conscious or unconscious inference to the best explanation (IBE) of why the speaker said what he did. According to IBE, one considers several possible explanations, and infers the truth of the explanation he deems best. Lipton (2007, p. 244) gives an example of a man who rings his doorbell and claims that Lipton's rain gutters are loose. To decide whether to trust him, Lipton constructs two theories – T_1 : the man is telling the truth; T_2 : he is lying and hoping to make a fast buck. Because Lipton deems T_2 a better explanation, he infers its truth, and does not trust the man. Social values, such as stereotypes, may affect the choice between T_1 and T_2 . If this is so, the role of values in affecting trust in testimony boils down to filling the underdetermination gap.

How strong is this objection? First, even on Thagard's and Lipton's accounts, IBE is not always involved in trustworthiness assessments. By default, a hearer accepts what she is told without engaging in conscious evaluation or inference, but there may be triggers, such as reasons for suspicion, which toggle the hearer to an evaluative mode (Lipton, 2007, pp. 240-241; Thagard, 2006, pp. 297-298). It is consistent with their accounts that social values influence which factors serve as triggers to begin with. For example, a speaker's being black may serve as a trigger for racist white hearers, but not for other people. Thus, Thagard's and Lipton's models allow values to influence trust *prior to* filling the underdetermination gap.

Second, that testimonial beliefs, even when subject to reflective processes, are inferential is controversial. Audi (1997) argues that testimonial beliefs are not inferential, even in cases going beyond default acceptance. Audi describes how a woman on a plane tells him that a philosopher lost his temper in a conference. At first, he suspends judgement, but as the story advances, he starts believing her, as she and the story seem more credible to him. At no place in the conversation, so Audi argues from introspection, has he engaged in inference. Rather, he has gradually realized a disposition to believe her, and come to trust her. Audi does not deny that the brain may be engaged in subconscious information processing, but not in inference, because inference entails belief formation, and at no time has he formed a belief *that* the woman is trustworthy.

Audi regards trust as a non-propositional attitude, namely, a *stance*, which means that trusting somebody and believing that he is trustworthy are two distinct mental states. Holton argues that being a stance distinguishes trust from mere reliance. A stance of trust entails a

this matter. In this case, he may be said to betray my epistemic trust, but arguably not my overall trust.

readiness to feel betrayal should it be disappointed, and gratitude should it be upheld. Holton's argument draws *inter alia* on his observation that unlike belief, trust, at least sometimes, is under our voluntary control. For example, in a popular exercise in drama class, when you let yourself fall back, there is a moment at which you voluntarily decide whether to trust your partner to catch you (1994, p. 63-67).

Lahno (2001) similarly argues that trust is an *emotional attitude*, which is a stance, rather than a belief, and the emotional nature of trust distinguishes it from mere reliance. Lahno argues that trust is a participant attitude in which a person regards herself and another as involved in interaction. Trust is characterized by the disposition to have certain emotions toward that person. Emotions affect the way we see the world. A person in love tends to see the world through pink glasses – he has the readiness to see everything positively, while a person in a gloomy state has an opposite readiness. Such readiness is not a belief about the world, but a stance toward it. Trust has a similar effect. A person whose good friend has been accused of a crime may trust her testimony despite the incriminating evidence because of his affection toward her (Lahno, 2001, p. 171-178). This emotional dimension of trust is missing from Lipton's and Thagard's inferential models.

Sociologists Lewis and Weigert argue that only a multifaceted conception of trust can account for its explanatory role in sociological theories of its manifestations – trust in testimony, in personal relations, in institutions, etc.⁵ Trust has “distinct cognitive, emotional, and behavioural dimensions which are merged into a unitary social experience” (1985, p. 969). They argue that game theorists' understanding of trust as a rational expectation is too narrow, as it looks only at its behavioural aspects overlooking its emotional dimensions (1985, pp. 974-978).

Summing the discussion so far, a strong case exists against the theoretical account of trust in testimony, or at least, against its being complete. At least sometimes or in some part, trust is a stance with an emotional dimension, rather than an inferential belief. A hearer may award trust to a speaker without conscious or unconscious choice of theory about the speaker's trustworthiness. In such cases, the role values play cannot be filling the underdetermination gap.

Can we explain how subjects extend trust without engaging in inference that involves theory choice? Fricker (2007, pp. 69-71) does so by drawing on virtue ethics and epistemology. She proposes a testimonial perceptual capacity she calls “testimonial sensibility”, which is a dispositional trait to react to testimony in certain ways in certain circumstances and form beliefs accordingly. When a subject's testimonial sensitivity functions optimally, i.e., when the subject correctly assesses speakers' trustworthiness, the subject is virtuous.

Fricker (2007, pp. 75-80) lists two features of testimonial sensibility, which illustrate why the effect of values on trust in testimony is not another aspect their gap-filling role. First, testimonial sensitivity has an emotional component. It entails emotions such as sympathy, suspicion, respect, or contempt, which reliably guide the virtuous person regarding whom to trust. Second, the process is not codifiable, namely, neither is it based on the application of a theory specifying when to trust, nor can it be formulated as a theory.

Testimonial sensibility is socially situated and affected by prevailing social values, but may be trained during a reflexive person's lifetime (Fricker, 2007, p. 82). Social psychology stresses the role of social stereotypes as heuristics subjects use to facilitate their credibility judgments. Some stereotypes are unreliable, e.g., that women are incapable of abstract

⁵ Lewis & Weigert's argument about the nature of trust from its explanatory role in sociology may persuade readers who are unimpressed by arguments from introspection, such as Audi's and Holton's.

thinking, and some are reliable, e.g., that second-hand car salesmen are dishonest about the cars they sell. When speakers hold stereotypes as prejudice, i.e., when they do not change them despite counterevidence, they are not virtuous (Fricker, 2007, pp. 31-42).

Fricker argues that “prejudice presents an obstacle to truth, either directly by causing the hearer to miss out on a particular truth, or indirectly by creating blockages in the circulation of critical ideas” (2007, p. 43). Systematic distortions in people’s credibility assessments are related to social power. The powerful are considered more trustworthy than they are, and the disempowered are considered less trustworthy than they are (Fricker, 2007, pp. 119-120). This makes the science of the day view those in power as authoritative and trustworthy and those without power as untrustworthy.⁶ Thus, even without filling the underdetermination gap, by systematically affecting trust in testimony, social values have serious epistemic ramifications.

I argued that social values affect people’s trust in testimony. There is a strong case for the view that testimonial beliefs, at least sometimes and in some part, are not inferential, and trust is a stance rather than a propositional attitude. Hence, social values play a different role from filling the underdetermination gap, and their affect may still have serious epistemic consequences. In the next section, I review psychological evidence about how values affect evidence assessment. Drawing on these studies, in sections 5 and 6, I argue that values may affect our assessments of the credibility of evidence itself.

3. Motivated Reasoning and Evidence Assessment

In the previous section, I argued that values affect the trust or credibility people accord to speakers’ testimonies. Sometime, however, the testifier’s identity is unknown or immaterial, but our assessment of it *qua* evidence is still similarly influenced by social values. For instance, Fricker (2007, pp. 34-35) describes blind referees for a journal who are prejudiced against a new methodology, rather than a person. They resist the evidence because of some countervailing motivational investment, such as loyalty to the old methods, or fear of intellectual innovation. In this section, I review research from experimental psychology that studies such effects of values on people’s evidential reasoning, on which I will draw later on.

Psychologists have studied the influence of people’s values, preferences, and incentives on their belief formation. “Motivated reasoning” denotes any process of reasoning that is affected by a person’s preference, wish or desire concerning the *outcome* of the reasoning process (Kunda, 1990, p. 480). Motivated reasoning is a species of confirmation bias, which is people’s tendency to form beliefs that reaffirm their prior beliefs and existing biases (Nickerson, 1998; Klayman, 1995).

Motivated reasoning affects evidence assessment. People assess the same evidence differently based on their directional goals. Here are some representative examples. Coffee lovers who read a scientific article claiming that caffeine was hazardous were less convinced by it than non-caffeine consumers. Sports fans were told that a previously winning team had lost a game. Fans of the team tended to see this as a mere fluke, while fans of the opposing team

⁶ For example, the U.S. Supreme Court Decision *Plessy* (1896), which stated that racial segregation was constitutional, reflected the biological theories of the time, according to which blacks were inferior to whites. These theories reflected blacks’ disempowered position and their lack of credibility as knowers (Southern, 1987, p. 147). Similarly, Darwin regards the ability to explain why women have lower intellectual capacities than men as a strength of his theory (1871, p. 326). This view reflects the social inequality between men and women in Victorian England. Many other examples are available. For racial theories of intelligence in the 19th and 20th centuries, see Gould (1996); for racial biological theories in the 19th and early 20th century, see Bowler & Morus (2005) at 415-437; for gender bias in biology in the 19th and 20th century see Okruhlik (1998); for gender bias in science of the Enlightenment see Bowler & Morus (2005) at 487-510.

tended to see this as a turning point (Kunda, 1990, pp. 488-490). Proponents and opponents of capital punishment received the same mixed evidence about its effectiveness. Both regarded the evidence as reaffirming their prior beliefs. When presented with the same studies, scientists tended to deem the studies that supported their previous beliefs as more methodologically sound than those that did not (Klayman, 1995, pp. 394-395).

Motivated reasoning is hard to eradicate. The success of debiasing techniques in experimental settings is modest at best. Even when subjects are trained to avoid certain biases, in a wide class of cases, they end up forming biased beliefs. Even successful debiasing techniques are not very effective in real-life conditions outside the laboratory (Lilienfeld et al., 2009).

The influence of directional goals on subjects' evidence assessment, though significant, is constrained. Motivated subjects are not at liberty to conclude whatever they want. They are constrained by their ability to rationalize their reasoning. Subjects "attempt to be rational and to construct a justification of their desired conclusion that would persuade a dispassionate observer. They draw the desired conclusion only if they can muster up the evidence necessary to support it" (Kunda, 1990, pp. 482-483). Subjects maintain an *illusion of objectivity*. They want to appear to themselves and others as following seemingly rational reasoning processes. This objectivity is illusory because they do not realize that they are biased, and that if they had different directional goals, they would probably form different or even opposite beliefs given the same evidence (Kunda, 1990, p. 483).

How is it possible for subjects to treat the same evidence differently? How can a smoker and a non-smoker, for example, treat the same evidence about the dangers of smoking differently? Psychologists identify a number of cognitive methods such as selective accessing of different memories on different occasions or choosing those reasoning methods that are likely to lead to the desired conclusion (Kunda 1990, pp. 486-489). But it is still puzzling from an *epistemic* perspective, rather than cognitive, how such differential treatment of the same evidence is possible.

One explanation, of course, is that the logical gap between theory and evidence allows values to enter and affect subjects' choice between rival hypotheses. However, this answer is partial, and cannot account for the variety of ways values are involved in evidential reasoning. In the next sections, I identify additional ways that enable people to treat the same evidence differently. Section 4 explains how values raise and lower the threshold of evidence required for justified epistemic judgments. This explanation applies to cases such as the caffeine consumers, smokers, and sport fans. Section 5 explains how values affect the relative weighing of discordant evidence. This explanation applies to cases such as the mixed evidence on capital punishment, and the scientists' methodology assessment.

4. Social Values Lower and Raise Evidential Thresholds

In this section I argue that values lower and raise evidential thresholds. They affect the strictness of threshold tests evidence is required to meet in a given context. The evidential threshold-adjusting role is different from the gap-filling role for two interrelated reasons. *First*, it occurs at a logically and often temporally prior step to filling the underdetermination gap. *Second*, in this role, values do not bridge, either top-down or bottom-up, the logical gap between evidence and theory. Namely, they constitute neither theoretical virtues, which constitute reasons for preferring one theory over another, nor reasons for adopting theoretical background assumptions.

I illustrate this with examples of evidential reasoning in science. The first example is distinguishing signal and noise in physics. I first argue that distinguishing signal from noise

may be influenced by social values. Then I argue that this influence is not captured by the gap-filling model because the influence occurs at a logically and temporally prior step to filling the underdetermination gap that is neither mediated by theoretical background assumptions nor guided by theoretical virtues.

In experimental physics, raw data is never pure. There is always some noise, which needs removing or reducing. The signal/noise distinction is not clear cut. Any dataset can be mathematically presented as a sum of a relatively simple and regular pattern and a certain level of noise. This means that any dataset can formally be described as the sum of one of infinitely many distinct patterns and a corresponding incidence of noise (McAllister, 1997, pp. 219-220).⁷ The same data may therefore lend itself to more than one partition between signal and noise. In practice, this may create a problem of distinguishing signal from noise. As Grinnell (1999, p. 207) writes:

In research at the edge of discovery, the difference between data and noise often is not obvious. Discovery at the forefront of knowledge requires learning to recognize something when one doesn't know beforehand what it looks like. Choosing what counts for data will depend on an investigator's experience and intuition – in short, his/her creative insight.

Since the right distinction is not always obvious, scientists may debate it. Social values affect this debate. Scientists are subject to social influences. They are under pressure to present their research in the strongest and most promising way in order to overcome scepticism, get funding, and publish. The psychological studies reviewed in the previous section suggest that scientists motivated toward a certain outcome are more likely see signal where others see noise, and vice versa.

To understand why the effect of values on distinguishing signal from noise is not a manifestation of the underdetermination gap-filling model, let us examine how physicists discriminate signal and noise. Philosophers observe that significant parts of this process are not theory-guided. Hacking writes the following about experimental physicists debugging a polarizing electron gun:

Debugging is not a matter of theoretically explaining or predicting what is going wrong. It is partly a matter of getting rid of “noise” in the apparatus. Although it also has a precise meaning, “noise” often means *all the events that are not understood by any theory*. The instrument must be able to isolate, physically, the properties of the entities that we wish to use, and damp down all the other effects that might get in our way (1983, p. 265; emphasis added).

Brown (1994, pp. 128-129) discusses noise reduction in a high-energy event in a bubble chamber. The left side of Figure 1 is a photograph of the event – the raw data, and the right side is a drawing of the event alone, with the noise removed. Brown (1994, p. 129) writes: “theories explain what is happening on the right; they never try to cope with the mess on the left”. Woodward (1989, p. 397) similarly notes that theories do not explain the data in its entirety, but rather the data after it has been analyzed, and the phenomenon of interest has been separated from extraneous background noise.

⁷ This is a consequence of the fact that any mathematical function f can be represented as a sum of two functions g and r , such that $f(x) = g(x) + r(x)$, g is a regular function, and r is the difference.

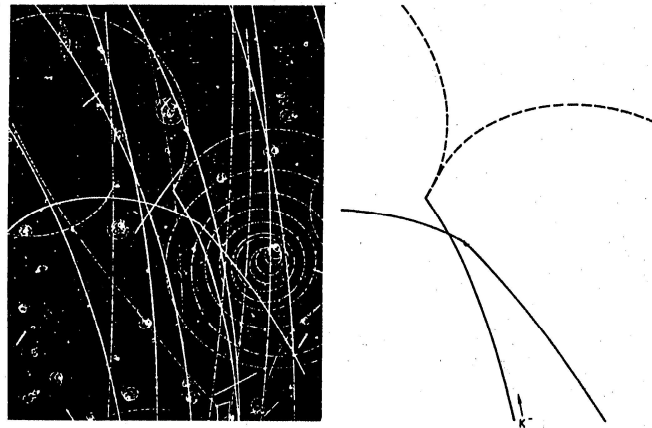


Figure 1: Signal and noise in a high-energy event. Source: Brown (1994, p. 128)

To clear noise, then, physicists often need not appeal to any theory to explain it away. As I will now argue, this means that the influence of values on distinguishing signal from noise cannot be another aspect of filling the underdetermination gap. In making this claim, I am drawing on the familiar distinction between raw data and the model of the data, which Frigg & Hartmann (2012) sum up as follows:

A model of data is a corrected, rectified, regimented, and in many instances idealized version of the data we gain from immediate observation, the so-called raw data. Characteristically, one first eliminates errors (e.g. removes points from the record that are due to faulty observation) and then present the data in a 'neat' way, for instance by drawing a smooth curve through a set of points.

As stated, theory explains the model of the data, not the raw data. Clearing noise, removing outliers, and the like are part of the process of constructing the model of the data from the raw data, which may be logically and temporally prior to theoretically explaining the data. As we have seen, classifying a datum as "signal" or "noise" is not necessarily backed up theoretically. Namely, scientists may classify a datum without theoretically justifying their decision. *Specifically, the scientists need not necessarily make any theoretical background assumptions to serve as reasons for the classification, and because there is no theory choice involved, they need not appeal to theoretical virtues to guide the process.* Recall, however, that values fill the underdetermination gap either bottom-up, by serving as reasons for adopting certain background assumptions, or top-down, as tie-breakers that represent theoretical virtues. It follows that the influence of values on the processes of distinguishing signal from noise may not amount to filling the underdetermination gap.

Rather than filling the underdetermination gap, values participate in determining the evidential threshold level putative evidence must meet. This is their second role according to WAVE. In the signal-noise example, values *inter alia* determine how clear and distinct a pattern in the data needs to be to count as signal, rather than being dismissed as noise. Sometimes, the data may not pass the threshold level, which means that it is all noise, which does not call for any theoretical explanation (I discuss such a case in Section 6.2).

The second example of how values participate in setting evidential thresholds is the choice of threshold values of statistical significance. Statistical studies use mathematical metrics to evaluate evidence. One such metric, known as significance level (α) or critical p -value, is, roughly speaking, the acceptable threshold for the probability of wrongly accepting a

false hypothesis. Another metric is *relative risk* (RR), which is the ratio of the rate of the occurrence of a condition in a group exposed to a putative risk factor and the respective rate in a non-exposed group.

There are commonly used numerical values for such metrics. A five percent significance level is commonly used for deciding whether to accept or reject a hypothesis. While such thresholds are sometimes treated as objective, invariant tests of significance, for example, in a number of U.S. courts' toxic torts decisions (Beecher-Monas, 2002, p. 64), this interpretation of their meaning is mistaken. Epidemiologists warn against a rigid and dogmatic application of threshold values. They argue that it is wrong to categorically accept studies whose significance level is lower than five percent and reject all others: "the actual p value should be reported and considered, not simply whether it falls above or below an arbitrary point" (Bryant & Reinert, 2001, p. S32).

Rather than being absolute indicators of significance, Hacking (1992, p. 152) regards such common statistical thresholds as "a technology of intersubjectivity", whose aim is indicating that a certain protocol was used and providing a method for intersubjective, interest comparisons within the scientific community. Wilholt (2009, p. 98) similarly characterizes them as *conventional* standards that impose implicit constraints on acceptable error probabilities within a research community. They are solutions to a social-epistemic problem of coordination in a community: They allow individual researchers to develop a reliable sense of the dependability of certain kinds of scientific outcomes based only on their knowledge of the procedures, rather than of the person who conducted the studies. Namely, provided that a scientist has adequately followed the relevant conventions, other scientists can reliably estimate the reliability of her reported outcomes without knowing her personally.

According to Wilholt, "the standards adopted are arbitrary in the sense that there could have been a different solution to the same coordination problem, but once a specific solution is socially adopted, it is in a certain sense binding" (2009, p. 98). These values, however, are arbitrary *only to a certain extent and within a certain range*. The conventional critical p -value could have been 6 or 4.6 percent. Such values would also have served as reasonable solutions to the community's coordination problem. A critical p -value of 45 percent, however, would not have worked, as it would have meant that the community accepted as statistically significant results that were just slightly higher than chance.

We can identify two levels of influence of social values on setting threshold values: the individual, and the community. At both levels, their influence is not restricted to filling the underdetermination gap. At the individual scientist's level, values, such as personal or ideological investments, may bias her judgment and cause her to infringe an explicit or implicit conventional standard by lowering or raising the evidential thresholds in a way that increases the likelihood of arriving at her preferred result, and violates her community's shared understanding of these thresholds (Wilholt, 2009, p. 99).

At the community level, values may influence the conventional threshold values themselves. Different social values in different scientific contexts may participate in raising or lowering conventional threshold values. Such change of communal conventions has nothing to do with filling the underdetermination gap. For example, in significance testing, there is an inherent mathematical trade-off between minimizing false positives and false negatives. Values influence the balance between false positives and false negatives. The existing scientific standards, which are manifested *inter alia* in the widespread choice of the five percent significance level, are conservative in that they regard false positives as more serious errors than false negatives (Wilholt, 2009, p. 99).

In some social contexts, scientists may adopt lower evidential thresholds, while in others they may adopt higher thresholds. In a society where smoking is customary and pervasive, scientists may set higher thresholds for evidence about the dangers of smoking than in a society that disapproves of smoking. *Ceteris paribus*, the same evidence for the dangers of smoking may be considered insufficient in a smoking-friendly society and sufficient in a smoking-disapproving society. This may happen without a need arising in the smoking-friendly society to come up with rival theories to explain away the evidence, as the underdetermination model requires. When social values and norms change, evidential thresholds may change accordingly. If it becomes less socially acceptable to smoke for whatever reasons, the evidential thresholds may drop accordingly, again without a need arising for any theoretical justification for this drop. Therefore, when we explain changes in prevailing theories and beliefs in a society over time, WAVE points our attention to the changing social values and their possible influence on the change in evidential thresholds.

It might be objected that setting statistical evidential thresholds can be captured by the logic of filling the underdetermination gap. Suppose that a dataset and a p -value of 0.05 yield theory T , and this same dataset with a p -value of 0.1 yield T' . This is a case of underdetermination; the dataset can be accommodated by more than one theory. Suppose also that values affect the setting of a p -value. It seems that in this case, values are playing a gap-filling role, or so this objection goes.⁸

In reply, I acknowledge that values fill the underdetermination gap in this case (supposedly, T and T' explain why the same dataset with different p -values yield different theories; otherwise, it is unclear in what sense the datasets “yield” them). But I argue that in other, more typical cases they do not. In typical cases, the choice is not between two theories, but between either accepting the test hypothesis, or rejecting it and accepting the null hypothesis. The null hypothesis is not theory in any meaningful sense. It does not explain the results, but merely states that the test hypothesis is false. Just like classifying signal and noise, setting a critical p -value need not be theoretically driven. Scientists need not necessarily make theoretical background assumptions to serve as reasons for setting a certain critical p -value rather than another, and since the null hypothesis is not a theory, it does not have theoretical virtues. The influence of values on setting the threshold value, then, need not be mediated or informed by theoretical considerations; hence, the influence of values on it need not amount to filling the underdetermination gap.

To further support this point, note that a value of $\alpha=0.05$, for example, is invariant under any specific theory. It holds just the same for a study about domestic violence among immigrant communities, hammer-throwing as serious leisure among middle-class women, and spatial orientation in cockroaches. Its choice is not driven by theoretical considerations, but pragmatic ones – it merely corresponds to the level of certainty a scientific community finds acceptable, rather than general or specific features of good theories.

5. Values Affect the Relative Weighing of Discordant Evidence

So far I identified two role values play in evidential reasoning according to WAVE. First, values affect trust in testimony. Second, they influence the evidential thresholds required for justified epistemic judgments, such as belief formation or theory acceptance. These roles are distinct from filling the underdetermination gap, because when we decide to trust a speaker or that the evidence falls short of an evidential threshold, we do not necessarily explicitly or implicitly consider rival theories that explain the evidence. Values serve as reasons for neither employing

⁸ I thank an anonymous reviewer for this objection.

theoretical background assumptions nor preferring one theory over another. Their role does not amount to filling the underdetermination gap either bottom-up or top-down.

In this section, I identify a third role: Values affect relative weighing of discordant evidence. In science, multimodal evidence, i.e., evidence from multiple techniques for the same theory, is often discordant. There are two types of discordance: inconsistency – an apparent contradiction between the hypotheses the evidence supports, and incongruity – different results that were produced under different background assumptions and are reported in “different languages” using different, possibly incommensurable units (Stegenga, 2009).

There is no algorithmic or universally agreed method to combine multimodal evidence. Qualitative methods, e.g., literature review, require judgment. Even quantitative methods, like meta-analysis, require judgment in their correct application at two stages: choosing the relevant evidence to begin with, and choosing the amalgamation method. Different inputs to the same method may produce different outputs. Different amalgamation methods, e.g., different meta-analysis methods, may also produce different outcomes for the same evidence (Miller, 2013, pp. 1310-1311; Douglas, 2012; Stegenga, 2011). Even proponents of evidence-based medicine, who advocate minimizing clinical judgment and basing decisions on methodically-amalgamated evidence, have not managed to eliminate judgment from clinical reasoning (Braude, 2012, Ch.6) or agree on a single evidence hierarchy. Rather, several evidence hierarchies exist, which all purport to implement the principles of evidence-based medicine. They weigh and rank different evidence differently, and may yield different conclusions for the same evidence (Upshur 2003).

In the face of multimodal discordant evidence, we may thus ask: Which evidence is more relevant to the given case? What relative weight should be given to each type of evidence? Which evidence should be most trusted?

Just like values lower or raise, to some extent, evidential thresholds, they also decrease or increase, to some extent, the relative weight an individual or a group assigns to different types of evidence within a body of multimodal discordant evidence. Given the same body of evidence, different persons or groups that adhere to different values may assign different relative weights to different evidence, without any need to theoretically justify the different weighing, i.e., without values filling the underdetermination gap.

An example from Miller (forthcoming a) illustrates this. A series of legal trials was held in U.S. Federal Courts about whether the drug Bendectin caused birth defects in human embryos. Among the multimodal discordant evidence were structural activity studies, in vitro studies, in vivo studies, and human epidemiological studies. Each evidence type has its merits and drawbacks (Miller, forthcoming a, Table 1). While human studies did not show a correlation between Bendectin and birth defects, a minority of in vitro studies (dose-response animal studies) showed a correlation, and the other evidence supported this possibility. The courts eventually ruled that human epidemiological studies were required for establishing causation in humans; namely, that epidemiological studies carried the bulk of the evidential weight in settling the issue. Hence, the plaintiffs did not prove that Bendectin caused birth defects.

Miller argues that this outcome was not inevitable, but was rather negotiated during litigation. How could social values influence this decision? The U.S. legal system is a jury system, where the jury – twelve citizens with no legal background – assume the role of the fact finder. The outcome of the Bendectin litigation was established when concerns were growing, especially among conservatives, about jurors’ abilities to perform their fact-finding role in trials involving complex scientific evidence. It was argued in popular and academic monographs that jurors tend to sympathize with plaintiffs (typically ordinary working-class

citizens) and be hostile to defendants (typically big corporations). Thus, they tended to decide for plaintiffs based on “junk science”. Such a practice by the jury, it was argued, not only averted justice, but also inhibited economic growth, as it compelled big corporations to spend major funds for legal defence and compensation.

In light of such public concerns, social values, namely, the disapproval of using the legal system to extort big corporations, might have influenced the courts toward assigning more weight to human epidemiological studies at the expense of other evidence. By contrast, in the regulatory context, where values of preventing health hazards prevail, human studies are *not* required for establishing causation. The International Agency for Research on Cancer (IARC) and the U.S. National Toxicology Program, for instance, have identified at least five human carcinogens and about 75 percent of probable human carcinogens based on *animal studies alone* (Cranor, 2005, p. 189). The different social values in the legal and regulative contexts is a plausible explanation of the different weighing of discordant evidence in similar evidential cases.

Just as the influence of social values on evidential thresholds is restricted, so is their influence on relative weighing of discordant evidence. The same psychological constraints of self-rationalization, and the constraints imposed by the need of the evidential standards to constitute effective solutions to communal coordination problems operate here too. While there may be more than one way to weigh the same multimodal discordant evidence, not any weighing will do. For example, in the Bendectin controversy, only a minority of animal studies showed a correlation. It is plausible that if *most* of the animal studies had shown a correlation, courts would have faced more difficulties dismissing them and assigning the bulk of weight to human studies. Nevertheless, as the comparison between the legal and regulatory contexts reveals, the impact of social values may still be significant.

WAVE’s third role may seem similar to Kuhn’s (1977) account of the top-down role of values in theory choice, discussed in Section 1.1. Let me clarify the differences between it and WAVE, and by doing so, also sharpen the differences between the third role and the gap-filling role.⁹ Kuhn argues that epistemic values need to be interpreted and weighed against one another. One interpretation and weighting scheme may lead to accepting one theory, while another may lead to accepting another. For example, suppose T is simpler, but has narrower scope than T' , then *ceteris paribus*, scientists who rank simplicity over scope will prefer to accept T , while scientists who rank scope over simplicity will prefer T' .

Kuhn’s account may seem similar to WAVE, but they are distinct. In Kuhn’s account, *values* are weighed against one another, whereas in WAVE, *evidence* is weighed against each other. Kuhn fails to distinguish weighing values from weighing evidence, which makes his analysis of rival theory choice lacking.¹⁰ Kuhn (1979, Ch.6&7) argues that in a crisis, anomalies accumulate, and scientists gradually lose faith in the current theory’s ability to explain them. This eventually leads them to adopt a new theory. To see why Kuhn’s account of values does not fully capture this dynamic, note that a paradigm shift is possible even without any change in the weighing of values. Suppose a scientific community vacillates between T_{old} and T_{new} , and

⁹ I thank an anonymous reviewer for pressing me on this point.

¹⁰ According to Dorn (2001), Kuhn was indeed not entirely satisfied with his explanation of when anomalies trigger a crisis and when they do not. Kuhn noted that usually anomalies are present during normal science, but count as puzzles that can be successfully solved either in the present or future. Kuhn has not managed to distinguish between anomalies that are crisis-evoking, and those that are not. By characterizing anomalies as counterevidence, and social values as factors that *inter alia* determine the relative weights that are assigned to them, WAVE constitutes an analytic framework that partly answers Kuhn’s challenge.

all members of the community agree on the same relative weighing of values and their application to the theories. Suppose that they agree that T_{new} fares as well as, or less than T_{old} regarding every epistemic value they share; namely, T_{new} is as or less simple as T_{old} , as or less broad in scope than T_{old} , etc.¹¹ If the scientists have been increasingly concerned about some anomalies and come to worry that T_{old} cannot overcome them, they may favour T_{new} . In this case, there has not been any change in the relative weighing of values. The only change has been in the relative weighing of evidence, i.e., in the relative weights the community assigns to the anomalies *qua* evidence against T_{old} within the overall available body of evidence. Only WAVE characterizes the logical relations between values and evidence in this example.

One might object that this example can be accommodated within Kuhn's account. According to this objection, the value the community most cherishes is fruitfulness, i.e., a theory's ability to "disclose new phenomena or previously unnoted relationships among those already known" (Kuhn, 1977, p. 322), and it deems T_{new} more fruitful than T_{old} . Even if we accept this objection, however, it does not threaten WAVE. The community deems T_{new} more fruitful *because* it has come to deem the anomalies more decisive evidence against T_{old} than before. The anomalies have been there all along; the change has occurred in the relative significance the community assigns to them *qua* evidence against T_{old} . The change in the relative weighing of discordant evidence according to WAVE, then, explains the change in the community's application of the value of fruitfulness to T_{old} according to Kuhn. Even if we interpret this example as a gap-filling case, the weight-adjusting role is still logically prior to, and distinct from the gap-filling role.

This concludes my presentation of WAVE, and the three roles values play in evidential reasoning, which are distinct from filling the underdetermination gap. In the next section, I illustrate how WAVE helps better analyze existing examples from science studies.

6. Using WAVE in Analyzing Empirical Case Studies

I identified three roles of values in evidential reasoning and argued that they are different from filling the logical gap of underdetermination. Existing case studies assume that the gap-filling model covers all epistemically relevant roles of values in evidential reasoning, or do not distinguish the gap-filling from other roles. By reanalyzing two case studies, I illustrate how WAVE can assist HPS/STS scholars in their research.

6.1. Dioxin Cancer Research

The first example is research on the carcinogenicity of dioxin, which Douglas uses to illustrate her argument from inductive risk (aka "the argument from error") against the model of value-free science. Douglas (2000; 2009, Ch.5) distinguishes two roles of values in evidential reasoning: direct and indirect. In their *direct* role, values serve as reasons for making an epistemic judgment. For Douglas, the direct role is illegitimate in the context of justification, as it amounts to wishful thinking. But in their *indirect* role, values determine the threshold that evidence must meet for making a justified judgment by determining tolerable levels of inductive risk. Douglas identifies two types of inductive risks: wrongly accepting a false hypothesis, and wrongly rejecting a true hypothesis. There is an inherent trade-off between them. Douglas argues that the indirect role is legitimate and required, because social values determine acceptable risks in a given context, and different social circumstances legitimately require different balances between types of errors. When we value a risk as mild, we lower the threshold required for making an evidential judgment, and when the risk is high, we raise it.

Clearly, Douglas' indirect role is the second role according to WAVE. Therefore, *Douglas' indirect role is not another aspect of filling the underdetermination gap*. Thus, Douglas' argument

¹¹ My example is highly idealized and is meant to make a purely logical point, rather than historical.

from inductive risk is immune to the usual objections to the underdetermination model discussed in Section 1.2.

Moreover, WAVE strengthens Douglas' case for the legitimacy of the indirect role. Douglas (2000) discusses a series of studies in which rats were exposed to dioxin, and slides with their tissues were taken to determine if they had developed cancer. Researchers needed to characterize the slides by identifying certain visual patterns in them. Three different studies that used *the same slides as data* characterized some slides differently. Douglas argues that values should not influence the characterization of clear evidence, namely, clear cases of diseased or healthy tissues, because this would amount to their playing a direct role. But values should play an indirect role by influencing the characterization of borderline evidence, by determining how clear a visual pattern needs to be for characterizing a slide as diseased. *Ceteris paribus*, in a society more concerned with the dangers of cancer, scientists should tend to characterize borderline slides as diseased, while in a society more concerned with the burdens of overregulation, the tendency should be characterizing them as healthy. This reflects the types and levels of inductive risk society is willing to take.¹²

It is still unclear, however, why the indirect role is permissible. Douglas argues that values should not influence characterizing clear slides because "our preferences for the world have no direct bearing on the way it actually is [...] If our empirical reasoning were guided by such wishful thinking, we would have little chance of allowing the world to surprise us" (2008, pp. 9-10). But doesn't the same argument apply to borderline evidence? A borderline slide, just like a clear slide, is either diseased or a healthy. We may characterize it correctly or incorrectly, but this will not change how the world actually is. Why, then, according to Douglas, is the indirect role epistemically legitimate in borderline cases?

WAVE may support a better argument for the indirect role. Researchers are most prone to motivated reasoning in borderline cases, because they are much more at liberty than in clear cases to rationalize a biased characterization. The less a visual pattern is clear, the more a researcher is prone to the influence of social values. Since an implicit influence of researchers' idiosyncratic values is already present and hard to avoid, we might as well make the influence of values on evidence characterization explicit and principled, recognize the values society deems important, and consciously consider them when we evaluate the evidence. Or so an argument that draws on WAVE for the legitimacy of the indirect role might run.

6.2. Gravity-Waves Research

WAVE sheds new light on the gravity-waves controversy. General Relativity predicts that moving massive bodies produce weak radiation-like phenomenon known as "gravity waves".

¹² The challenge to value-free science from the indirect role of values was already posed by Rudner (1953), who argued that scientists cannot avoid weighing social risks against each other when deciding to accept or reject hypotheses. A standard reply to Rudner was that the need to evaluate research consequences arises only at the last stage of inquiry in the context of application; thus, scientists need not be those who weigh social risks (Jeffery 1966; McMullin 1983). This reply accords with the received model of value-free science. For example, Hempel (1965) agrees that social values play an indirect role, but assumes a strict division of labour: Science assigns various hypotheses probabilities, while society assigns them utilities and decides which hypotheses to accept. A novel feature of Douglas' work is showing that this reply is inadequate, because a strict separation between basic and applied science, or the contexts of justification and application, is unsustainable. The dioxin example illustrates that values penetrate deep into the context of justification, and affect various stages of research, such as study design, data analysis, evidence characterization, and evidence interpretation. When research outcomes finally reach the context of application, they are already saturated with social value judgments, and reflect the various trade-offs between values that were made in the process of inquiry leading to them.

Collins (1981b; 1985; 2004) followed a controversy in the physicists community in the 1970's over gravity waves detection. Gravity waves are hard to detect because isolating their effects from other forms of radiation is very difficult. Joseph Weber claimed to have detected gravity waves. While in the early 1970's Weber's claim enjoyed some credibility, toward the end of the decade, a consensus emerged that his claim was incorrect. Weber's critics argued that his detected values were significantly higher than theory predicted, that there were difficulties in replication, that his computer data analysis was flawed, and that he had found a correlation between two isolated detectors which were later discovered to be asynchronous, and hence could not possibly have detected the same events (Collins, 1981b, pp. 38-44; 2004, Ch.9).

Nevertheless, Collins argues, evidence alone could not settle the controversy. Social values were also required. Collins (1981a) endorses the underdetermination thesis, and argues that sociological studies of science have shown underdetermination to be an actual phenomenon of day-to-day science, rather than a mere philosophical abstraction. He thus adopts a threefold methodology for analyzing controversies and their closure: First, identify the theories that are underdetermined by the evidence; second, identify the local and global social values that operate and the theories associated with them; third, identify the values responsible for the closure (cf. Finn, 2011, p. 84).

Collins argues that Weber had seemingly rational replies to all the criticism and that none of his colleagues found *all* of the criticism persuasive (2004, Ch.10; 1981b, pp. 49-54). In response, Franklin (1994) argues that physicists' rejection of Weber's claims was reasoned and rational. The dispute between Collins and Franklin boils down to whether social values were required for bridging the underdetermination gap or whether rational reasons sufficed. This debate misses important aspects of the case. The controversy centred on the evidential threshold required for a justified detection claim, namely, whether Weber successfully built a gravity-wave detector, or whether his detector was producing "pure noise". As WAVE teaches us, values may lower and raise evidential thresholds *without* bridging the underdetermination gap, i.e., without scientists having a theoretical justification for rejecting Weber's detection claims. Values might also influence Weber's perceived credibility among his colleagues without their providing any justification for it. This is what actually happened later, when physicists practically ignored or dismissed Weber's testimonies, including a 1977 paper in *Nature*, without theoretically engaging with them at all (Collins, 2004, pp. 201-205).

Almassi (2009) argues that Weber and his critics were simultaneously rational in their contrary beliefs, which illustrates that rational disagreement between epistemic peers is possible. Collins makes a similar claim (1994, p. 502). If this were a reasonable disagreement where both sides were rational, *any* outcome of the dispute might arguably be rational. This seems to be Collins' own view, as he argues that although the outcome of communal disputes is contingent on social values, in esoteric fields, such as gravity-waves, the research community has an epistemic right to settle them by forming a consensus, and lay outsiders should defer to the consensus when they seek justified belief on the disputed matter (Collins & Evans, 2002, pp. 242-243; Koerth-Baker, 2011).

WAVE militates against interpreting this affair as a reasonable peer disagreement. Motivated reasoning casts doubt on Weber's reasoning objectively and rationally about his own data, especially when the data was borderline between signal and noise. Thus, the two sides were not equally reasonable. In other cases, the majority that forms the consensus may suffer from biases. Thus, pace Collins, WAVE shows that without further analysis of the social dynamics of the case, deference to a consensus may be a bad idea (cf. Miller 2013; forthcoming a).

WAVE thus takes us beyond the limited prism of the underdetermination gap-filling model. Analysis of case studies involving evidential reasoning should closely attend to the influence of social values on trust in testimony, evidential thresholds, and relative weighing of discordant evidence. It must examine how values change over time or between contexts, and how they lead to different outcomes in similar evidential contexts.

Conclusion

The Weight-Adjusting Account of Values and Evidence (WAVE) identifies three roles social values play in evidential justification, besides their familiar role of filling the logical gap of underdetermination of theory by evidence. First, values affect trust in testimony. Second, values lower and raise evidential thresholds. Third, values affect the relative weighing of multimodal discordant evidence.

Critics of the underdetermination gap-filling model have been wrong to assume that the model covers all epistemically relevant influences of values on evidential reasoning. They have too quickly concluded that their criticism entails that the effect of social values on scientific evidential reasoning is scarce and modest. Research in experimental psychology shows that the effect of values on evidential reasoning, while constrained, is nevertheless not trivial. I illustrated that these effects are not fully captured by the underdetermination model, which needs to be complemented by WAVE. Science studies scholars who wish to account for differences in beliefs or theories in different periods or different social contexts should thus look into the influences that WAVE identifies in their case-study analysis.

References

- Almassi, B. (2009). Conflicting expert testimony and the search for gravitational waves. *Philosophy of Science*, 76(5), 570-584.
- Audi, R. (1997). The place of testimony in the fabric of knowledge and justification. *American Philosophical Quarterly*, 34(4), 405-422.
- Baier, A. (1986). Trust and antitrust. *Ethics*, 96(2), 231-260.
- Beecher-Monas, E. (2002). *Evaluating scientific evidence: An interdisciplinary Framework for intellectual due process*. Cambridge: Cambridge University Press.
- Biddle, J. (2013). State of the field: Transient underdetermination and values in science. *Studies in History and Philosophy of Science*, 44(1), 124-133.
- Bowler, P. J., & Morus, I. R. (2005). *Making modern science: A historical survey*. Chicago: University of Chicago Press.
- Braude, H. D. (2012). *Intuition in medicine: A philosophical defense of clinical reasoning*. Chicago: University of Chicago Press.
- Brown, J. R. (1994). *Smoke and mirrors: How science reflects reality*. London: Routledge.
- Bryant, A. H., & Reinert, A. (2001). Epidemiology in the legal arena and the search for truth. *American Journal of Epidemiology*, 154(12), S27-S35.
- Carrier, M. (2008). The aim and structure of methodological theory. In L. Soler, H. Sankey, & P. Hoyningen-Huene (Eds.), *Rethinking scientific change and theory comparison* (pp. 273-290). Dordrecht: Springer.
- Carrier, M. (2011). Underdetermination as an epistemological test tube: Expounding hidden values of the scientific community. *Synthese*, 180(2), 189-204.

- Collins, H. (1981a). Stages in the empirical programme of relativism. *Social Studies of Science*, 11(1), 3-10.
- Collins, H. (1981b). Son of seven sexes: The social destruction of a physical phenomenon. *Social Studies of Science*, 11(1), 33-62.
- Collins, H. (1985). *Changing order: Replication and induction in scientific practice*. Chicago: University of Chicago Press.
- Collins, H. (1994). A strong confirmation of the experimenters' regress. *Studies in the History and Philosophy of Science*, 25(3), 493-503.
- Collins, H. (2004). *Gravity's shadow: The search for gravitational waves*. Chicago: University of Chicago Press.
- Collins, H. & Evans, R. (2002). The third wave of science studies: Studies of expertise and experience. *Social Studies of Science*, 32(2), 235-296.
- Cranor, C. E. (2005). The science veil over tort law policy: How should scientific evidence be utilized in toxic tort law? *Law and Philosophy* 24: 139-210.
- Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London: John Murray.
- Dorn, H. (2001). Kuhnian rainedance. *London Review of Books*, 23(17). <http://www.lrb.co.uk/v23/n17/letters#letter7>
- Douglas, H. (2000). Inductive risk and values in science. *Philosophy of Science*, 67(4), 559-579.
- Douglas, H. (2008). The role of values in expert reasoning. *Public Affairs Quarterly*, 22(1), 1-18.
- Douglas, H. (2009). *Science, policy, and the value-free ideal*. Pittsburgh, PA: University of Pittsburgh Press.
- Douglas, H. (2012). Weighing complex evidence in a democratic society. *The Kennedy Institute of Ethics Journal*, 22(2), 139-162.
- Duhem, P. (1954). *The aim and structure of physical theory*. Princeton: Princeton University Press.
- Elliot, K. (2011). *Is a little pollution good for you? Incorporating societal values in environmental research*. Oxford: Oxford University Press.
- Finn, C. (2011). *Science studies as naturalized philosophy*. Dordrecht: Springer.
- Franklin, A. (1994). How to avoid the experimenters' regress. *Studies in History and Philosophy of Science*, 25(1), 463-491.
- Fricker, E. (1995). Telling and trusting: Reductionism and anti-reductionism in the epistemology of testimony. *Mind*, 101(414), 393-411.
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford: Oxford University Press.
- Frigg, R., & Hartmann, S. (2012). Models in science. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, Fall 2012 Edition. <http://plato.stanford.edu/archives/fall2012/entries/models-science/>
- Gould, S. J. (1996). *The mismeasure of man*, rev ed. New York: Norton.

- Grinnell, F. (1999). Ambiguity, trust, and the responsible conduct of research. *Science and Engineering Ethics*, 5(2), 205-214.
- Haack, S. (1998). *Manifesto of a passionate moderate: Unfashionable essays*. Chicago: University of Chicago Press.
- Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press.
- Hacking, I. (1992). Statistical language, statistical truth and statistical reason: The self-authentication of a style of scientific reasoning. In E. McMullin (Ed.), *The social dimensions of science* (pp. 130-157). Notre Dame, IN: University of Notre Dame Press.
- Harding, S. (Ed.). 1976. *Can theories be refuted? Essays on the Duhem-Quine thesis*. Dordrecht: Springer.
- Hempel, C. (1965). Science and human values. In *Aspects of scientific explanation: And other essays in the philosophy of science* (pp. 81-96). New York: Free Press.
- Holton, R. (1994). Deciding to trust, coming to believe. *Australasian Journal of Philosophy*, 72(1), 63-76.
- Intemann, K. (2005). Feminism, underdetermination, and values in science. *Philosophy of Science*, 72(5), 1001-1012.
- Jeffrey, R. C. (1956). Valuation and acceptance of scientific hypotheses. *Philosophy of Science*, 23(3), 237-246.
- Kitcher, P. (2012). *Preludes to pragmatism: Toward a reconstruction of philosophy*. Oxford: Oxford University Press.
- Klayman, J. (1995). Varieties of confirmation bias. *Psychology of Learning and Motivation*, 32, 385-418
- Koerth-Baker, M. (2011). The scientist who studies scientists—An interview with Harry Collins. *Boing Boing* (Apr 14). <http://boingboing.net/2011/04/14/the-scientist-who-st.html>
- Kourany, J. (2010). *Philosophy of science after feminism*. Oxford: Oxford University Press.
- Kuhn, T. S. (1970). *The structure of scientific revolutions*, 2nd ed. Chicago: University of Chicago Press.
- Kuhn, T. S. (1977). Objectivity, value judgment, and theory choice. In *The essential tension: Selected studies in scientific tradition and change* (pp. 320-339). Chicago: University of Chicago Press.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480-498.
- Lahno, B. (2001). On the emotional character of trust. *Ethical Theory and Moral Practice*, 4, 171-189.
- Laudan, L. (2004). The epistemic, the cognitive, and the social. In P. Machamer & G. Wolters (Eds.), *Science, values, and objectivity*. Pittsburgh: University of Pittsburgh Press, 14-23.
- Laudan, L., & Leplin, J. (1991). Empirical equivalence and underdetermination. *The Journal of Philosophy*, 88(9), 449-472.
- Lewis, D. J., & Weigert, A. (1985) Trust as a social reality. *Social Forces*, 63(4), 967-985.

- Lilienfeld, S. O., Ammirati, R. and Lilienfeld, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives on Psychological Science*, 4(4), 390-398.
- Lipton, P. (2007). Alien abduction: Inference to the best explanation and the management of testimony. *Episteme*, 4(3), 238-251.
- Longino, H. (1990). *Science as Social Knowledge: Values and objectivity in Scientific Inquiry*. Princeton: Princeton University Press.
- Longino, H. (1995). Gender, politics and theoretical virtues. *Synthese*, 104, 383-397.
- Longino, H. (2002). *The fate of knowledge*. Princeton: Princeton University Press.
- Longino, H. (2004). How values can be good for science. In P. Machamer & G. Wolters (Eds.), *Science, Values and Objectivity* (pp. 127-141). Pittsburgh, PA: University of Pittsburgh Press.
- Machamer, P., & Douglas, H. (1999). Cognitive and social values. *Science & Education*, 8, 45-54.
- McAllister, J. (1997). Phenomena and patterns in data sets. *Erkenntnis*, 47, 217-228.
- McMullin, E. (1983). Values in science. In P. D. Asquith & T. Nickles (Eds.), *PSA 1982* (Vol. 2, pp. 3-28). East Lansing, MI: PSA.
- Miller, B. (2013). When is consensus knowledge based? Distinguishing shared knowledge from mere agreement. *Synthese*, 190(7), 1293-1316.
- Miller, B. (forthcoming a). Scientific consensus and expert testimony in courts: Lessons from the Bendectin litigation. *Foundations of Science*.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175-220.
- Norton, J. (2008). Must evidence underdetermine theory? In M. Carrier, D. A. Howard, & J. Kourany (Eds.), *The challenge of the social and the pressure of practice: Science and values revisited* (pp. 17-44). Pittsburgh: University of Pittsburgh Press
- Okruhlik, K. (1998). Gender and the biological sciences. In M. Curd and J. A. Cover (Eds.), *Philosophy of science: The central issues* (pp. 192-207). New York: Norton.
- Pinnick, C. (1994). Feminist epistemology: Implications for the philosophy of science. *Philosophy of Science*, 61(4), 646-657
- Plessy v. Ferguson. (1896). 163 U.S. 537.
- Potter, E. (1996). Underdetermination undeterred. In L. H. Nelson and J. Nelson (Eds.), *Feminism, Science, and the Philosophy of Science* (pp. 121-138). Dordrecht: Kluwer.
- Putnam, H. (2002). *The collapse of the fact/value dichotomy and other essays*. Cambridge, MA: Harvard University Press.
- Quine, W. V. O. (1951). Two dogmas of empiricism. *The Philosophical Review*, 60, 20-43.
- Rolin, K. (2004). Why gender is a relevant factor in the social epistemology of scientific inquiry. *Philosophy of Science*, 71(5), 880-891.
- Rudner, R. (1953). The scientist *qua* a scientist makes value judgments. *Philosophy of Science*, 20(1), 1-6.

- Solomon, M. (2001). *Social empiricism*. Cambridge, MA: MIT Press.
- Southern, D. W. (1987). *Gunnar Myrdal and black-white relations: The use and abuse of an American dilemma, 1944-1969*. Baton Rouge: Louisiana State University Press.
- Stanford, K. (2010). *Exceeding our grasp: Science, history, and the problem of unconceived alternatives*. Oxford: Oxford University Press.
- Stegenga, J. (2009). Robustness, discordance, and relevance. *Philosophy of Science*, 76(5), 650-661.
- Stegenga, J. (2011). Is meta-analysis the platinum standard of evidence? *Studies in History and Philosophy of Biological and Biomedical Science*, 42(4), 497-507.
- Taylor, C. (1992). *Sources of the self: The making of modern identity*. Cambridge, MA: Harvard University Press.
- Thagard, P. (2006). Testimony, credibility, and explanatory coherence. *Erkenntnis*, 63, 295-316.
- Upshur, R. E. G. (2003). Are all evidence-based practices alike? Problems in the ranking of evidence. *Canadian Medical Association Journal*, 169(7), 672-673.
- Wilholt, T. (2009). Bias and values in scientific research. *Studies in History and Philosophy of Science*, 40(1), 92-101.
- Woodward, J. (1989). Data and phenomena. *Synthese*, 79, 393-472.