The Error Is in the Gap: Synthesizing Accounts for Societal Values in Science

Christopher ChoGlueck*†

Kevin Elliott and others separate two common arguments for the legitimacy of societal values in scientific reasoning as the *gap* and the *error* arguments (respectively, the arguments from underdetermination and from inductive risk). This article poses two questions: How are these two arguments related, and what can we learn from their interrelation? I contend that we can better understand the error argument as nested within the gap because the error is a limited case of the gap with narrower features. Furthermore, this nestedness provides philosophers with conceptual tools for analyzing more robustly how values pervade science.

1. Introduction. Increasingly, philosophers have rejected value-free ideals of science and turned their attention to examining values in concrete cases and developing alternative norms for legitimate/illegitimate influences (see Hicks 2014).[1] This article argues that such attempts would benefit ana-

Received May 2017; revised February 2018.

*To contact the author, please write to: Department of History and Philosophy of Science and Medicine, 1165 East Third Street, Morrison Hall 302, Bloomington, IN 47405; e-mail: cglueck@iu.edu.

†I am grateful for advice and encouragement from Lisa Lloyd and Jutta Schickore; for thought-provoking discussions with Josua Aponte-Serrano, Evan Arnet, Jordi Cat, Kevin Elliott, Kate Grauvogel, Amit Hagar, Nora Hangel, Bennett Holman, Becca Jackson, Ryan Ketcham, Monica Morrison, and Nick Zautra; and for critical comments from two anonymous reviewers. An earlier version of this article was presented at Indiana University for the 2017 Hanson Prize Lecture, and I thank the audience for their questions and criticism. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under grant no. 1342962. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

1. Following the discussion among philosophers of science, I use "value" to mean a good or desirable quality, such as the accuracy of a theory or the fairness of a procedure (see McMullin 1982, 5). In contrast to rules that determine choices algorithmically, val-

Philosophy of Science, 85 (October 2018) pp. 704–725. 0031-8248/2018/8504-0008$10.00
Copyright 2018 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/699191 Published online by Cambridge University Press

lytically from a more comprehensive understanding of value-ladenness by returning to the original critiques of value-freedom.

The value-free ideal of science narrows the role for social, ethical, and political values—taken to be distinct from scientific, epistemic, and cognitive values—in scientific reasoning and practice (Douglas 2009; Elliott 2011). Defenders of this value-freedom accept the legitimacy of social, ethical, and political values only in the early and late stages of science, such as with funding and technological applications. The ideal proscribes the use of these purportedly nonscientific values within the so-called internal core of scientific reasoning, especially in evaluating evidential support for a hypothesis (i.e., theory choice).

If, however, scientific reasoning ought to embody the value-free ideal, then it must be the case that scientific reasoning could be free from these extrascientific values. Elliott has attempted to cast doubt on the achievability of this ideal with two lines of reasoning, which he distinguishes as the *gap* and *error* arguments (2011, 62). Both arguments, he claims, entail at times the ineliminability of societal values, which he defines as "qualities or states of affairs that societies or social groups hold to be good or desirable" (59).[2] The gap argument, for one, posits a "logical gap or underdetermination between theory and evidence," requiring something to connect evidence with theory (62). The beliefs or assumptions that play this connective role may involve societal values because "epistemic or constitutive values are insufficient to determine which background beliefs should be accepted" (62).[3] He distinguishes this gap line of reasoning from the error argument, which focuses on the risks one is willing to take in accepting or rejecting a hypothesis. The level of acceptable *inductive risk* is based on one's valuation of the consequences that might result from an erroneous choice. Thus, when a potential error has societal consequences, the amount of evidence needed to accept or reject a hypothesis is partially an ethical, political, or social question.[4]

Since Elliott introduced this framing in 2011, several philosophers have used the gap/error distinction (or underdetermination/inductive risk) for their philosophical projects (e.g., Betz 2013; Brown 2013; Brigandt 2015). More recently, a handbook in the philosophy of science and a reference book on ethics in science and engineering now present the gap and error arguments

ues merely guide judgments (Kuhn 1977), e.g., the appraisal of means (evaluation) and the prizing of ends (valuing; Dewey 1939, 5; McMullin 1982, 5).

2. Elliott's *societal* values mirrors Longino's (1990) *contextual* values, which she contrasts with the *constitutive* values of scientific communities.

3. The term "the gap argument" comes from Intemann (2005). Elliott cites Don Howard, Janet Kourany, and Helen Longino as proponents of the gap line of argumentation.

4. As error proponents, Elliott cites Heather Douglas, C. West Churchman, Richard Rudner, Kristin Shrader-Frechette, Phillip Kitcher, and Carl Cranor.

https://doi.org/10.1086/699191 Published online by Cambridge University Press

separately to more general audiences in their "Values in Science" entries (Biddle 2015; Douglas 2016). While these arguments appear quite different, how distinct are they?

Despite its attractiveness, Elliot's gap/error distinction has the potential to obscure several philosophically significant similarities between the two. Defending the ideal of value-free science, Betz (2013, n. 7) criticizes Elliott's separation of the two arguments because both posit an inferential gap between evidence and theory; therefore, he merges the two into one argument. Biddle (2013, n. 3) likewise claims that the inductive risk argument is a "special case" of underdetermination, although he does not discuss his reasoning. Brown has argued that, despite their differences, the gap and error arguments share a premise about evidence and uncertainty. Because they "begin from a situation where the evidence is fixed and take values to play a role in the space that is left over," both assume that the positive role for values is secured by the limitations of evidence (Brown 2013, 834). Elliott himself maintains that the two arguments rely on three common principles, namely, scientists' ethical responsibilities, the underdetermination of conclusions by evidence, and the social harms of value-neutrality (2011, 80). These claims of similarity suggest the further question: how precisely are the gap and error arguments related?

Beyond merely noting their shared characteristics, I contend that the two arguments are better understood as *nested*—the error in the gap—because the error argument is a special, limited case of gap. Sections 2 and 3 analyze them into their constitutive parts, which I will call *features* rather than premises to emphasize that these arguments are more like *analytical frameworks* for investigating the relationship between evidence, values, and theories.[5] Because each argument comes from several sources, this requires me to interpret and present a single cogent version of my own understanding.[6] While I am committed to the veracity of both arguments as presented here, my goal is not to defend them from objections but to discuss their interrelation.[7] Accordingly, section 4 explains how the features of the error argument "fit" into those of the gap.

It is not my intention to reduce the error to the gap but to synthesize them for further insight into the relationship between values and science. Together they provide a comprehensive and cohesive account of how values pervade

5. I thank an anonymous reviewer for suggesting this shift.

6. I am interpreting the arguments and presenting them in my own terms and in their strongest possible form. Thus, I am striving primarily for strength and clarity of the arguments rather than fidelity to the various textual presentations of them.

7. For further criticism and defense of the gap argument, see Giere (2003), Anderson (2004), Okruhlik (2004), Intemann (2005), Howard (2006), Carrier (2011), and Goldenberg (2015). For criticism and defense of the error argument, see Churchman (1956), Jeffrey (1956), Levi (1960), Wilholt (2009), Steele (2012), Betz (2013), Biddle (2016), and de Melo-Martin and Intemann (2016).

https://doi.org/10.1086/699191 Published online by Cambridge University Press

science along a scale of empirical uncertainties, from abstract to concrete, with respective value judgments determining the multifaceted ways that theory is related to observation. Section 5 articulates this framework to refute Winsberg's (2012) claim that pervasiveness of values "in the nooks and crannies" of climate models disallows philosophical analysis. Instead, the synthetic framework suggested by this nestedness can improve how we investigate value-ladenness in concrete cases. But before we can understand this nestedness, we first need to examine them individually. Let us start with the gap.

2. The Gap Argument. The gap (or underdetermination) argument against value-free science goes as follows. Within the framework of empiricism, underdetermination requires a bridge to span the gap between observations and theory. Science operates within a larger context, so its results often have societal stakes. Rather than striving for unachievable detachment, incorporation of alternative values can improve the objectivity of science by promoting intersubjectivity through interaction and standardization. Thus, to improve empirical research, evidential assessments ought to allow rather than exclude societal values.

This section draws on the work of Helen Longino along with Otto Neurath and Janet Kourany to describe how empirical uncertainty creates the space for societal values to improve scientific knowledge.[8] The gap argument involves four features (G1–G4) abbreviated as empirical uncertainty, a bridge feature, societal stakes of scientific knowledge, and objectivity through intersubjectivity. I will argue that these features form a cohesive critique of the value-free ideal, and then I will provide an example from evolutionary biology to illustrate the gap framework in action, particularly how feminist values can improve science via critique.

2.1. Empirical Uncertainty (G1). The gap argument is situated within a larger empiricist framework, including how scientists make decisions (Neurath), reason about evidence (Longino), and choose between theories (Kourany). Following Neurath, take the Cartesian problem for whoever "wants to create a world-view or a scientific system" of making rational, practical decisions based on limited empirical knowledge and "doubtful premises" (Neurath 1913/1983, 3). In science, this problem is a procedural one that occurs at the community level. What are the procedures, asks Longino (1990), by which communities of scientists reason empirically about evidence? How are they, inquires Kourany (2003a, 2003b), to make decisions about planning their research?

8. I prioritize Longino's account because it is the strongest and most developed version. Howard (2006) pulls his gap argument primarily from Neurath, so I have chosen to use the latter as my primary proponent (see also Uebel 2005, nn. 10, 11; Brown 2013; pace Okruhlik 2004). Note that Kourany (2010) no longer appeals to underdetermination.

https://doi.org/10.1086/699191 Published online by Cambridge University Press

Scientists here are faced with the problem of underdetermination: given an empirical basis for knowledge, without assuming foundational certainty, observations can only go so far. The empirical uncertainty, as Neurath notes, is bidirectional. For one, observations are theory laden, so their veracity cannot be considered individually or independent of theory. Moreover, every attempt at theory revision is contingent on prior ideas and experience, so the reviser is left without any suggestions for altering concepts. Thus, how is one to decide empirically "on the basis of insufficient insight" (Neurath 1913/1983, 4; see also Cartwright et al. 1996)? For evidential reasoning more specifically, Longino argues that there is a "logical gap" between the state of affairs that is said to be evidence and the hypothesis for which that state of affairs is said to be evidence (1990, 52). One might, for instance, take a child's red-spotted stomach as evidence that she has measles. However, the connection is not self-evident from the observation of her spots, nor does the hypothesis itself say why red spots are evidence of measles. Why take an observation as evidence of a hypothesis? What determines the "evidential relevance" of an observation to a hypothesis, filling in the gap between the two (43)?

2.2. A Bridge Feature (G2). To overcome this gap, each of these accounts supplies a bridge feature. For Neurath, the "auxiliary motive" is what provides an "aid to the vacillating" by producing practical certainty from empirical and logical uncertainty (1913/1983, 4). The "purest form" of the auxiliary motive is casting lots, while alternative means include following tradition, casting votes, and maximizing simplicity (5–8; see Cartwright et al. 1996; Okruhlik 2004; Reisch 2005, 31). Longino's "background assumptions" accomplish a similar function for providing reasons to accept a hypothesis based on certain observations (1990, 43). Why should red spots be taken as evidence for measles? Perhaps red spots are a symptom of the disease, or perhaps the doctor says so, or perhaps the crystal ball reader says so. Kourany (2003a, 10) ascribes the bridging function to the goals, values, and ideals of the scientific community, including those from the broader society.

Thus, the bridge features that enable scientists to reason empirically come from additional theoretical commitments, often derived from the scientific community. But, what does this bridge look like, and what does it have to do with societal values? That is, even if empirical uncertainty appears to "makes space" for societal values, it remains unresolved whether societal values can and ought to operate as bridges in science. Features G3 and G4 provide this further justification legitimizing societal values prima facie.

2.3. Societal Stakes of Scientific Knowledge (G3). Because of the social position of scientific communities, the gap proponents argue that scientific reasoning often cannot be neutral to political or social aims. Because

https://doi.org/10.1086/699191 Published online by Cambridge University Press

science and society are interrelated, there is a mutualistic exchange of values between them, allowing scientific communities to incorporate societal values, particularly in the form of bridge features. For this reason, Longino (1990, 1996) argues that the background assumptions held by scientists are often based on contextual/societal values or have "political valence" in certain social contexts, disallowing a sharp distinction between scientific and societal values.

Furthermore, scientific communities are authoritative producers of knowledge. Their epistemic authority can create societal stakes for the scientific enterprise, supporting or frustrating political agendas. Neurath discusses the potential for science to be an ally to the proletariat. The institutional establishment of Catholicism, capitalism, and fascism in Austria were all buttressed by the metaphysics of the day, which obscured the economic and social situation of the working class. Thus, by jettisoning speculative metaphysics, Neurath believed that materialist science provided the path to proletariat liberation (1928/1973, 297; see also Cat, Chang, and Cartwright 1991; Howard 2003). Kourany and Longino, on the flip side, discuss the harms perpetuated by the scientific community against women, such as the legacy in psychology and biology of explaining and justifying the inferiority of women to men (Kourany 2003a, 3; for a more extensive discussion, see Longino 1990, 103–32; Kourany 2010, 3–19). For instance, in Kourany's rejoinder to Ronald Giere's suggestion to "withhold judgment" rather than going "beyond the data" in cases of underdetermination (Giere 2003, 20), she claims that such inaction "is frequently not feasible in real scientific practice, at least the real scientific practice that concerns us here" (Kourany 2003b, 24). In these politicized contexts, the political neutrality of science is not possible because of its authority and other countervailing political forces; therefore, the scientific community is either for or against social progress. For the gap argument, this entails that how scientists choose to bridge the gap in these cases has unavoidable societal stakes because it either supports or challenges societal interests.

2.4. Objectivity through Intersubjectivity (G4). When societal stakes are unavoidable, societal values inevitably play the role of a bridge feature. Therefore, the gap proponents appeal to community-level processes to improve how scientific communities overcome underdetermination. More specifically, they allow societal values to play a legitimate role as bridge features if they promote objectivity (in the intersubjective sense) through critical interaction or standardizing procedures. These intersubjective senses of objectivity contrast *detached objectivity*, where individual scientists seek to distance themselves from any interest other than truth and knowledge (see Douglas 2009, 122; Lloyd and Schweizer 2014, 2068). Kuhn (1977), McMullin (1982), and Laudan (1984), for instance, argue that underdetermi-

https://doi.org/10.1086/699191 Published online by Cambridge University Press

nation implies a logical gap between evidence and theory, but they reserve the bridging role to scientific, epistemic, and cognitive values alone. For Kuhn, McMullin, and Laudan societal bridge features are biases in the negative sense of a value's distorting force on the quest for objective truth. In contrast, the gap proponents consider the presence of societal values at the individual level to be an inevitable form of subtle or unconscious bias that is not necessarily tainting. In fact, it is the very presence of background assumptions and auxiliary motives that allows individual subjects to collect and evaluate observations as evidentially relevant (Longino 1990) and to make decisions empirically despite insufficient insight (Neurath 1913/1983).

Thus, the common feature between Longino and Neurath is their rejection of detached objectivity, which enables their shift to the communal level for objectivity through interaction or standardization.[9] The illegitimate form of bias for the gap proponents exists more prominently at the community level when the societal values bridging the gap go unquestioned. For instance, when most of a scientific community holds the same assumptions, the group neglects certain perspectives with their respective background assumptions, ignoring relevant evidence and compromising peer evaluation. Longino's *interactive objectivity*, which seeks to preserve the reliability of science, is constituted socially via critical engagement (Longino 1990; Douglas 2009, 128; Lloyd and Schweizer 2014, 2070). Moreover, because bad bias is a community-level problem of epistemic homogeneity, this form of intersubjective objectivity makes for the possibility of good bias at the individual level by exposing and scrutinizing taken-for-granted bridge features. Thus, as an "oppositional science," feminist critics can improve scientific rigor by critiquing "mainstream" interpretive frameworks and attempting to provide alternative explanations (Longino 1990, 214).

Along with making interactions more rigorous, societal values can also promote objectivity between subjects by consolidating which sorts of bridge features are legitimate. Whereas Longino focuses on critical interactions, Neurath is more concerned with *procedural objectivity* by setting social standards to promote understanding and collaboration (Douglas 2009, 125–26; Lloyd and Schweizer 2014, 2069). Neurath argues that scientific communities need to coordinate knowledge at the community level for action, so they should develop a universally comprehensible language of science (Cat et al. 1991). Marxist physicalism enables such intersubjective, transdisciplinary collaboration. Rather than relying on metaphysical abstractions such as "the spirit of the people," Marxists emphasize concrete institutions and behaviors and thus improve social and economic sciences (Neurath 1928/1973, 293; see also Howard 2003). Thus, because they promote intersubjectivity Marxist values would benefit the sciences by stan-

9. I thank an anonymous reviewer for pressing me to clarify this feature.

https://doi.org/10.1086/699191 Published online by Cambridge University Press

dardizing how scientists bridge the gap (i.e., in concrete, material terms). The potential for objectivity and epistemic improvement via these forms of intersubjectivity provides the normative grounds for why the gap proponents allow societal values in scientific reasoning, at least prima facie (see also Harding 1991; Hicks 2014).

To recapitulate, the gap argument follows from empirical uncertainty (G1), where a bridge feature (G2) is needed to span the gap between observation and theory. Because of the social authority of scientific communities, knowledge production does not occur in a social vacuum without political stakes (G3). Moreover, because of the potential epistemic improvements from incorporating societal values into the social processes of science (G4), these values can be legitimate bridge features. The gap argument undermines the value-free ideal by emphasizing the need for nondetached forms of objectivity because of empirical uncertainty and the context of knowledge production.

2.5. Feminist Criticism in Evolutionary Biology. A clear case from Lloyd's work demonstrates how the gap argument works in practice. Lloyd (2005) has explored different evolutionary explanations for female orgasm in humans. From a feminist perspective, she critiques much of the biological research as androcentric. Because male orgasm results from intercourse and thus increases men's evolutionary fitness, some biologists have simply assumed that female orgasm also results from intercourse, likewise increasing women's fitness. On the basis of that evidence about male sexuality and the additional bridging assumption that male and female sexualities are the same, many biologists have argued for the adaptive hypothesis that female orgasm increases fitness and thus was the result of selective pressure on our female ancestors.

As an external feminist critic, Lloyd criticized this explanation as supported more by androcentric bias widespread in the evolutionary-biology community than by empirical evidence. As she points out, sexology research has documented that females do not reliably orgasm with intercourse. However, the androcentric assumption of the similarity between males and females obscured these observations and disappeared the distinctness of female sexuality. Note how societal values (i.e., androcentrism and feminism) influenced what scientists assumed and what they counted/ignored as evidence. Impelled by a feminist perspective to investigate the evidential reasoning of this adaptive hypothesis, Lloyd (2005) argued that the androcentric bridge feature (that male sexuality provides a model for female sexuality) was undermined by relevant but ignored data.

In addition, notice how certain feminist societal values that had been excluded from the community improved this biological research via critical interaction. The gap argument does not reduce scientific disputes into value

https://doi.org/10.1086/699191 Published online by Cambridge University Press

conflicts. Rather, it demonstrates how societal values can enhance a community's evidential reasoning, such as when the debate over Lloyd's critique prompted biologists to measure the fitness of female orgasm, and they found no correlation (Zietsch and Santtila 2013). This finding provided disconfirming evidence for the androcentric bridge feature (that male and female sexualities are the same) that had gone unquestioned previously.

Thus, recognizing community-level bias and correcting its epistemic problems often requires external critiques based on alternative societal values (see Jukola [2015] for the case of external criticism of biomedicine's bias toward commercial interests). Lloyd's study exemplifies how including alternative societal values in the process of evaluating unquestioned bridge features has in fact improved the products of science by making the process more objective in the interactive sense. With this analysis under our belt, let us now turn to the error argument, which philosophers of science have typically contrasted with the gap.

3. The Error Argument. The error (or inductive risk) argument is another prominent line of reasoning that challenges the value-free ideal. It goes as follows: given evidential limitations, the level of evidence sufficient to make a reliable inference is uncertain. Thus, scientists need standards of evidence to set the burden of proof for them to accept a hypothesis. However, different standards have different rates of scientific error, which can result in different risks in terms of ethical and social consequences. Because one must choose a standard that favors certain errors over others, including errors with social impact, scientists ought to consider the potential social consequences of their methodological choices.

This section marshals the work of Heather Douglas, as well as C. West Churchman, Richard Rudner, Carl Hempel, Carl Cranor, and Kristin Shrader-Frechette, to clarify how the potential for error entails the need for ethical considerations in scientific reasoning.[10] The error argument also has four features (E1–E4) abbreviated as evidential uncertainty, standards of evidence, social consequences of scientific error, and responsibility to consider social consequences. After describing how these features challenge the value-free ideal, I will illustrate the error argument with an example from risk assessment in governmental regulation.

3.1. Evidential Uncertainty (E1). The error argument begins within the context of a scientist faced with a set of evidence for a given hypothesis or theory. Because one's evidence is always limited, there is potential for an erroneous inference. The error proponents have various descriptions for the

10. I have chosen to prioritize Douglas's account because it is the most thoroughly developed and the most recent.

https://doi.org/10.1086/699191 Published online by Cambridge University Press

uncertainty inherent in making scientific inferences, including the problem of evidential limitations (Churchman 1948, 248), incomplete verification (Rudner 1953, 2), inductive risk (Hempel 1965, 92), and an inductive gap (Douglas 2009, 96). Consider again the diagnosis of measles. Granting that red spots are evidence of measles, how much evidence is sufficient to accept the hypothesis that a given child has measles? Are external red spots sufficient, and if so how many? Or would we also need to see Koplik spots inside his mouth? The issue here is the problem of induction: how much evidence is enough to make a general claim from knowledge of particulars?

3.2. Standards of Evidence (E2). To handle this evidential uncertainty, the error proponents contend that scientists need standards for determining the quality and quantity of evidence that is sufficient or adequate for adjudicating the veracity of a hypothesis (see Churchman 1948, 21; Rudner 1953, 2; Hempel 1965, 92; Douglas 2009, 103). The classic version of this sort of judgment involves hypothesis testing and the choice of a statistical threshold, such as statistical significance. Churchman discusses the "adequacy of a set of observations" for making an inference after an experiment (1948, 21). Rudner (1953, 2) asks whether "the evidence is sufficiently strong or . . . the probability is sufficiently high to warrant the acceptance of the hypothesis." More recently, Douglas (2009) has extended the argument to include other aspects of scientific judgment, such as the collection, characterization, interpretation, and presentation of data. She argues that just like the choice of a level for statistical significance, scientists need standards for characterizing data, such as distinguishing between borderline cases of malignant and benign tumors in rat kidneys (Douglas 2000).

3.3. Social Consequences of Scientific Error (E3). The choice of a methodological standard to overcome evidential uncertainty can have ethically significant consequences in society. Standards carry certain risks of scientific error, such as systematic over- and underestimation. With these scientific risks comes a "second-class of mistakes" when scientific errors have potential social consequences, such as the result of scientific knowledge for manufacturing goods, making policy, and adjudicating court cases (Cranor 1993, 13). The possibility for these second-class mistakes follows more from the social context of research than from researchers' intentions.

Nonetheless, valuation of these potential errors (both scientific and ethical/social) is based on one's interests. As Churchman (1948) argues, the very notion of error depends on one's goals and criteria for success. Those who aim to protect public interests, for instance, are more willing to risk overregulation, which places an increased burden on the industrial producer. In testing error rates, one might demand, as Rudner (1953) suggests, a higher burden of proof for vaccine toxicity compared to belt-buckle failure, such as

https://doi.org/10.1086/699191 Published online by Cambridge University Press

a greater number of tests or more experimental control. The justification of a standard follows from the alleged gravity of the potential ethical consequences, for example, embarrassed customers versus poisoned patients (Rudner 1953; see also Hempel 1965, 93). Thus, in cases with potential social fallout, the choice of a standard of evidence promotes certain societally relevant errors over others, regardless of the researchers' intent.

3.4. Responsibility to Consider Social Consequences (E4). Because of this coupling of scientific and social risks, which results from the context of knowledge production, scientists ought to consider both sorts of risk that follow from their methodology. The gap proponents thus acknowledge the impossibility of detachment from values and instead suggest making implicit moral valuations explicit. Thus, this accountability employs the *public* sense of objectivity, that judgments be made open to inspection and reflection (Lloyd and Schweizer 2014, 2068). Douglas bases this scientific responsibility on general moral responsibility and the absence of any special exemption for scientists (Douglas 2009, 66–86; see also Rudner 1953, 4; Churchman 1956; Hempel 1965, 93; Cranor 1993, 152–78; Shrader-Frechette and McCoy 1993, 192). The target of this proposed responsibility is not so much holding individuals accountable as legitimizing ethical reflection by scientists with political power, such as science advisers. Societal interests are already implicit in their methods (E3), so scientists ought to consider these values explicitly by contemplating the risks of their choices.

Note that this does not mean scientists are responsible for whatever happens because of their research. This concern is levied by Jeffrey (1956), who admonishes against Rudner's ascription of social significance to any particular scientific judgments because researchers have limited knowledge of the future uses of their results. How, for instance, can researchers choose a standard of evidence for vaccine toxicity when they do not know whether the vaccine would be for human children or pet monkeys? Granting that scientists are not omniscient, the responsibility incurred by the error argument must be limited to what scientists can reasonably foresee (Douglas 2009, 66–86). Especially for policy-oriented science, the applications and potential consequences are more clear.

To summarize, the error argument involves the following reasoning. It is uncertain how much evidence one must require to accept/reject a hypothesis (E1), so scientists need standards for assessing evidential sufficiency (E2). However, choosing a standard hinges on how serious an error would be, and this is an ethical question when there are potential social consequences (E3). Because scientists are responsible for the consequences they can reasonably foresee, they ought to account for broader social risks in addition to scientific ones in their choice of standards (E4). The error argument undermines the value-free ideal by demonstrating how societal values play a nec-

https://doi.org/10.1086/699191 Published online by Cambridge University Press

essary role in setting standards of evidence when social consequences are possible.

3.5. Regulatory Consequences from Risk Assessment. The classic example of the error argument comes from risk assessments in toxicology and ecology (Cranor 1993; Shrader-Frechette and McCoy 1993). Because scientists cannot assess harms directly, they must instead assess risks as estimates of the probability of harm. Such estimates are by their nature a fallible form of evidence, allowing for two possible scientific errors: false positives (mistakenly identifying a safe substance as harmful) and false negatives (mistakenly identifying a harmful substance as safe). In statistical terminology, these correspond to type I and type II errors, which can be reduced by increasing statistical significance and statistical power, respectively. These statistical risks incur two additional potential social consequences for risk managers: overregulation and underregulation. The risk of false positives varies inversely with that of false negatives, so scientists much choose one sort of statistical (and corresponding regulatory) error over the other. Normative assumptions about management and regulation, therefore, are an inevitable part of risk assessment methodology.

In the United States, for instance, to bring drugs to market, pharmaceutical companies must show them to be safe in phase 1 trials before assessing their efficacy.[11] This procedure is intended to minimize false negatives, and it places increased financial and evidential burdens on producers to conduct additional trials in the public's interest.[12] Note again that the error argument pinpoints the ineliminable role of societal values in certain social contexts when choosing standards of evidence, not supplanting evidence but enabling scientists to make inferences from limited evidence. Now, with both arguments analyzed into their constituents, let us examine how they are logically related.

4. The Error Is in the Gap. Together, the error and the gap arguments provide us with a forceful case for legitimizing societal values in science. While I find both arguments compelling, I will not defend them further here because my focus is on their interrelation (see n. 7). Building on these separate analyses, this section demonstrates how we can understand the error argument as a special, limited case of the gap. I argue that these two frameworks as presented in sections 2 and 3 rely on similar assumptions that share a nested relation. Nested assumptions are narrowed features involving

11. See http://www.fda.gov/ForPatients/Approvals/Drugs/ucm405622.htm.

12. Nonetheless, common practices, such as operationalizing harms as discrete outcomes (which excludes subtle, chronic ones) and not publishing phase 1 trials (especially for failures), systematically underestimate the harms of pharmaceuticals (Stegenga 2016).

https://doi.org/10.1086/699191 Published online by Cambridge University Press

type-token or concrete-abstract relations, such as examples or particular cases of more general claims. A detailed analysis of how each of the error's features (E1–E4) fits within those of the gap (G1–G4) in terms of their nested assumptions will clarify this relationship.

Take the first feature of the error argument: evidential uncertainty about how much evidence is enough (E1) is a case of the more general problem of empirical uncertainty regarding what counts as empirical evidence (G1). That is, the problem of induction that preoccupies the error proponents is one of the problems of underdetermination. Stanford (2016) distinguishes between two sorts of underdetermination—contrastive and holist. They differ by whether one holds fixed data or theory, and they illustrate how evidential uncertainty (E1) relates to empirical uncertainty (G1). In Stanford's *contrastive* underdetermination, one begins with a set of data and then reasons that multiple theoretic inferences are possible. Kourany appeals to this contrastive sort of underdetermination, which focuses on inferences from data to theory. There is a more limited sort of contrastive underdetermination that I call *enumerative* underdetermination, which focuses on the threshold at which one can make an inference from data to theory. This enumerative sort of contrastive underdetermination is the error argument's evidential uncertainty (E1), which focuses on questions of "how much?"

Stanford's *holist* underdetermination, however, posits that given a theory, whether and when to revise that theory is underdetermined by incoming observations. Longino concentrates on this direction of how theoretic assumptions influence what observations even count as evidence (G1). Holist underdetermination entails that the very category of empirical evidence is empirically underdetermined. Thus, the gap argument accounts for both holist and contrastive underdetermination—Neurath even discusses this bidirectionality. However, the error proponents' evidential uncertainty (E1) as the enumerative sort of contrastive underdetermination is a case of the more general class of empirical uncertainty (G1).

The second feature of the error argument, the standards of evidence (E2), is a paradigmatic example of a bridge feature (G2), which connects observations to theory as supportive evidence. For one, the methodological function of a standard of evidence (e.g., setting a statistical significance level for rejecting the null hypothesis) is a decision rule that bridges the gap between a limited set of data points and a hypothesis (Shrader-Frechette and McCoy 1993, 84, 157). Furthermore, by comparing their logical functions in evidential reasoning, one can see the nested relation. Determining evidential relevance (G2) is prior to deciding evidential sufficiency (E2). In Churchman's terms, before deciding "How many observations should he make?" an experimentalist must first decide "What are observations?" (1948, 21). Or, in Longino's terms, before she can ask "How much evidence e is sufficient for accepting hypothesis h?" a scientist must first ask "Is e even rele-

https://doi.org/10.1086/699191 Published online by Cambridge University Press

vant to *h* in the first place?" For instance, if one considers rodent data irrelevant to understanding human health, then no amount of rodent data will suffice for evidencing claims about humans.

Third, both arguments appeal to the larger social context to challenge claims of scientific neutrality because science shapes and is shaped by society (E3-G3). Nonetheless, the claim that ethically relevant consequences may follow from scientific errors (E3) is one case of how scientific knowledge production has societal stakes (G3), where the stakes are concrete risks rather than more abstract struggles over representation. The error proponents, on the one hand, focus on how preferences for over- and underestimation have implicit ethical valuations in certain contexts because of different rates of social error. On the other, the gap proponents extend their scope of consideration beyond error rates to more general and abstract aspects of contextual relevance. Longino (1996), for instance, argues that values of seemingly pure cognitive or scientific merit can have implicit societal valuations, which affect whether women are visible or not in scientific theories and thus can legitimate and reinforce certain political beliefs and agendas (see also Rooney 1992). In Lloyd's (2005) example, the social stakes include the visibility of a distinctly female sexuality in evolutionary accounts. Thus, the stakes of scientific knowledge production included in the gap's robust contextualism (G3) go beyond immediate, concrete consequences in society (E3) to the larger social and political forces supported and reinstated by scientific theories and practices.

The final features of each argument (G4 and E4) share the same neutral notion of individual bias as ineliminable and enabling by rejecting the detached sense of objectivity. Unlike other epistemologies of science, which conceptualize societal values as negative factors that distort the scientific pursuit of truth (e.g., McMullin 1982), the error and gap arguments rely on a neutral account of individual bias, such as in statistics. Statistical bias is the result of methodological choices, predisposing one's study to a certain conclusion. Such bias exists at abstract levels of methodology, such as how the operationalization of a concept can reduce the number of observations or how lumping two populations together can obscure their differences, and is conceptualized in the gap as background assumptions and auxiliary motives (G3). Bias also exists at more concrete levels, such as the choice of sample size, *p*-value, or rat-tumor key, which are the methodological choices of standards in the error argument (E3). The ineliminable presence of such biases from methodology justifies the rejection of the possibility of detachment.

Furthermore, these two normative features also conceptualize societal values as one of the key legitimate grounds for scrutinizing and evaluating claims about evidence. In the gap, explicitly including societal values into social processes can expose values' effects and improve how the gap is bridged (G4), and in the error, social responsibility is leveraged to make value

https://doi.org/10.1086/699191 Published online by Cambridge University Press

judgments public and hold scientists accountable (E4). Both approaches use societal values as grounds for (indirect) arbitration of evidence claims. Arguably, this common theory of evidence contrasts the "values as evidence" approach, which seeks to distinguish legitimate from illegitimate societal values by using evidence itself as the empirical arbiter of values (see Goldenberg 2015).[13]

Despite rejecting detached objectivity and legitimizing societal values in arbitrating evidence, E4 relies on a narrower sense of objectivity than G4. With objectivity through intersubjectivity rather than individual impartiality (G4), the gap argument appeals to community-level processes to improve interaction, correct hegemonic bias, and coordinate with standards. The error argument, likewise, by holding scientists responsible to consider the foreseeable social consequences of their methodology (E4) encourages making value judgments explicit and reflecting on them critically. Promoting reflexivity hinges on the public sense of objectivity, which enables accountability and interaction (Lloyd and Schweizer 2014, 2068). Nevertheless, this argument does not provide the additional normative force of the gap argument that the epistemic products of science will improve. Whereas objectivity through intersubjectivity (G4) justifies the scientific community's active cultivation of societal values because of their potential to aid empirical inquiry, the responsibility of scientists to consider social consequences (E4) is more suggestive and cautious in pointing toward possible ethical, rather than epistemic, improvements. Thus, while rejecting detachment and embracing societal values' relevance to evidence, the error argument's notion of objectivity (i.e., mere publicity rather than further interaction and standardization) is more limited than that of the gap.

Thus, all features of the error argument are nested within those of the gap. Accordingly, one could think of the error argument as a concrete example (or set of examples) of the gap, with the problem of induction as a particular sort of empirical uncertainty. In this case, choosing a standard of evidence to bridge the enumerative sort of contrastive underdetermination requires scientists to appraise societal risks, which are the social stakes of this methodological decision. Because of the authority of science in society and ineliminability of individual biases from statistical methodology, scientists ought to consider the social impacts their methodological choices might have.

One objection to my thesis that the error is nested in the gap is that the two arguments rely on different conceptions of scientific inference with different notions of validation. The error argument is premised on the deductive model of consequence testing, whereby one deduces the consequences of a theory's being true or false as hypotheses and then performs a decisive

13. I thank an anonymous reviewer for noting this commonality and contrast.

https://doi.org/10.1086/699191 Published online by Cambridge University Press

test to settle the theory's validity.[14] In contrast, the gap argument employs an inductive/abductive model of confirmation, whereby one looks at a variety of evidence and assesses validity based on coherence and robustness.[15] To use Nancy Cartwright's distinction between *clinching* and *vouching*, while error proponents conceptualize good evidence as categorically clinching a claim as true/false, gap proponents see it as merely vouching for relative likelihood or plausibility (see Cartwright 2011; Osimani and Mignini 2015). The objector might grant that the gap and the error have converged over the problem of theory assessment and the resulting uncertainties facing scientific inferences, be they deductive, inductive, or abductive. However, with different logics of validation (either clinching or vouching), the two arguments are more like alternative formulations.

In response, the deductive model of consequence testing is arguably a narrower form of validity than the abductive/inductive one of confirmation, thus allowing a version of the former to fit into the latter. Confirmational validity admits the importance of experimental tests when assessing robustness. While it does not regard them as clinching, confirmational validity does acknowledge decisive tests as a potentially strong voucher, especially given additional background knowledge. In contrast, consequence-testing validity ignores the evidential import of robust vouchers, focusing instead on decisive clinchers that simply surpass the standard of evidence (see Cartwright 2011).

Furthermore, and more to the point, this objection is based on a misunderstanding of my thesis. *Nesting* in this case does not entail reducing the error to the gap or eliminating the error argument as uninteresting or useless. That is, I am not claiming that the error simply *is* the gap. On the contrary, by nesting the two, I aim to preserve the integrity of both arguments by putting them in conversation. As I will argue in section 5, understanding them as nested allows us to synthesize their insights and better analyze the ways values pervade science along a scale of uncertainties from abstract to concrete.

To conclude, the novelty of this analysis is my claim that the gap and error arguments are not merely similar but closely interrelated. While the gap argument is general enough to accommodate the error as a limited case, the error is reliant on more narrow features. Nonetheless, they share many assumptions and provide a common analytic framework. This coheres with but expands on Elliott's analysis, wherein he notes that both arguments rely on three common principles: "(1) that scientists have ethical responsibilities to consider the major societal ramifications of their work; (2) that the evi-

14. The error argument's apparent reliance on deductive logic has been one locus of criticism (e.g., Jeffrey 1956).

15. Both Neurath and Longino have written extensively on holism and coherentism in science (see Cartwright et al. 1996; Longino 1990).

https://doi.org/10.1086/699191 Published online by Cambridge University Press

dence available to them is frequently not sufficient to determine their conclusions; and (3) that it is often socially harmful or impractical for scientists to withhold judgment or to provide uninterpreted, allegedly value-free information to decision makers" (2011, 80). These correspond respectively to features E4, G1-E1, and G3-E3. By making their parts more explicit and their relationship clearer, this analysis provides a fuller understanding of how the error fits within the gap.

5. Analyzing How Values Pervade the Gaps and Errors. *Why does it matter if the error is in the gap?* This nestedness has practical implications for philosophers by providing conceptual tools for analyzing more robustly how values permeate science. Most philosophers who specialize in values and science have abandoned the value-free ideal (pace Betz 2013), shifting to examining values in concrete cases and developing value-laden replacements (Hicks 2014; e.g., Anderson 2004; Douglas 2009; Kourany 2010; Elliott and McKaughan 2014; Intemann 2015; Brown and Havstad 2017). A key difficulty for progress in such investigations is scrutinizing values that are pervasive but obscured.[16]

One extreme stance of pessimism about uncovering value judgments in science comes from Winsberg (2012), who argues that the values ubiquitous "in the nooks and crannies" of climate models are unanalyzable (130). Climate modeling involves many uncertainties, including *structural model uncertainty* about the physics and mathematics used for conceptualizing and coding and *parameter uncertainty* about model inputs. While modelers have tried to manage these uncertainties with sampling and ensemble methods, Winsberg claims that such "model choices have reflected balances of inductive risk, and models have been optimized, over their history, to particular purposes, and to particular metrics of success" (124). Nonetheless, it is impossible to recover societal value judgments about structuring models and selecting parameters "in bite-sized pieces" such as "predictive preferences and inductive risks" because they "are buried in the historical past under the complexity, epistemic distributiveness, and generative entrenchment of climate models" (131). Furthermore, the inaccessibility of these past value judgments renders futile attempts to achieve transparency, accountability, and alignment with the "right values."

While Winsberg's claims about concealed values are certainly true, it is not clear that such obscurity from complexity, distribution, and entrenchment entails defeatism rather than the need for more careful examination. Furthermore, it ignores the nested framework of conceptual tools we have ready-made for such investigation. For lack of space, let us focus on how

16. I thank an anonymous reviewer for advising me to expound on the significance of this nestedness.

https://doi.org/10.1086/699191 Published online by Cambridge University Press

specific empirical uncertainties (G1-E1) correlate to bridge features (G2-E2) and how the latter interact across levels. Recall that we have understood the gap and error arguments as analytical frameworks, useful for investigating the manifold relationships between evidence, values, and theories. Nested together, the gap and the error arguments illustrate how values, including societal values, saturate scientific knowledge because of the need for decisions about different uncertainties (G1-E1). Since value judgments are necessary components of its constitution, they are "encoded" into scientific knowledge (Longino 1990, 216; Douglas 2009, 18)—not independently but rather as a multiplicity of interlocking bridge features for shoring up certainty, coherency, and confidence (G2-E2).

If viewed as separate, the gap and error arguments provide compatible but alternative approaches for understanding values at independent categories of uncertainty. However, once synthesized as a nested framework, we can see them instead as working at different levels of analysis and conceptualize their uncertainties as falling across logically interdependent levels along a scale of abstract to concrete. These uncertainties range from the holist underdetermination of the category of empirical evidence (G1) to the enumerative underdetermination of evidential thresholds for statistical inference (E1). Thus, determining the range of ways that theory relates to observation requires bridge features at each level, for example, providing evidential relevance (G2) and sufficiency (E2). Take the most abstract form of gap, how background assumptions help scientists constitute their object of inquiry. The preliminary characterization of research objects involves researchers' desires because it "depends not on what nature tells us but on what we wish to know about it" (Longino 1990, 99). Constituting the research object narrows which questions are appropriate and what responses register as answers. At a lower level of abstraction, value-laden background assumptions determine what evidence is relevant to the object of inquiry, guiding scientists toward collecting/measuring certain phenomena and ignoring others. Moving to the more concrete uncertainties associated with errors, valuations of risk determine whether the relevant evidence collected about the object of inquiry is sufficient to infer a given theory. The risk of error thus prompts choices about standards of evidence for drawing conclusions (Douglas 2009). Note, moreover, how this nestedness entails that there is interactivity across levels, where bridges at one level rely on and reinforce those at others, akin to what Elliott (2017, 166–68) calls a "tapestry of values." For instance, the sort of evidence one demands as relevant contributes to judgments about whether current evidence is enough, especially if that category of evidence is scarce or absent.

Responding to Winsberg's charge, we can propose for philosophical investigation several bridge features coupled with the uncertainties of climate modeling. Moving from abstract to concrete, what follows are examples of

https://doi.org/10.1086/699191 Published online by Cambridge University Press

nested levels (L1–L3) of uncertainty that required modelers to make decisions relevant to the societal values of improving globalized versus localized planning:

> (L1) *Constituting the object of inquiry* (for structural model uncertainty): How do we design the causal space of climate models? For global predictions, models need to account for interactions between atmosphere and ocean. However, regional modelers instead prioritize finer-grained details, for example, soil heterogeneity and species-specific evapotranspiration (Morrison 2018).
>
> (L2) *Evidential relevance* (for parameter uncertainty): What observational data are appropriate for confirming/calibrating these models? For global changes, we might look for upward trends in average surface temperature. For local concerns, we might instead focus on extreme regional events, such as floods and hurricanes (Intemann 2015).
>
> (L3) *Evidential sufficiency* (for uncertainty about standards): How much evidence of change is enough to attribute extreme weather events to climate change? Scientists often use frequentist statistics, preferring to assume no effect rather than risk positing one erroneously. However, to minimize forecast error and avoid the risk of underpreparation for catastrophes, we might want to use a Bayesian approach that assumes climate change is influencing every weather event and assesses how much (Mann, Lloyd, and Oreskes 2017).

Between these nested levels there are interactions connecting the various levels of design: the choices we make about causal space (L1) could influence model resolution and our ability to partition anthropogenic climate forcing for an extreme event (L3). Moreover, our desires to minimize the risks of underattribution and underpreparation (L3) have prompted some to conceptualize the causal space differently (L1), for example, prioritizing thermodynamic changes over other dynamical ones (Lloyd and Oreskes 2018). Contra Winsberg, the values associated with these levels are not hopelessly obscured "in the nooks and crannies" of the models but rather open to investigation as judgments associated with underdetermination gaps including, but not restricted to, inductive risks. Moreover, they are not atomized "in bite-sized pieces" at the microlevel of code but part of a broader framework, with interaction and coherence across interdependent levels.

6. Conclusion. By analyzing the error and gap arguments into their constituent features, this article has advanced the novel thesis that the error is a special case of the gap. This nestedness does not entail that the error

https://doi.org/10.1086/699191 Published online by Cambridge University Press

argument is trivial or that values are so pervasive that they are unanalyzable. Instead, it provides us with a set of conceptual tools to examine more comprehensively the full scale of uncertainties in science and their respective value judgments. Accordingly, we are better equipped to understand the gaps and errors throughout science.

REFERENCES

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.

Betz, Gregor. 2013. "In Defence of the Value Free Ideal." *European Journal for Philosophy of Science* 3 (2): 207–20.

Biddle, Justin. 2013. "State of the Field: Transient Underdetermination and Values in Science." *Studies in History and Philosophy of Science* A 44 (1): 124–33.

———. 2015. "Values in Science." In *Ethics, Science, Technology, and Engineering*, ed. J. Britt Holbrook. Farmington Hills, MI: Macmillan Reference.

———. 2016. "Inductive Risk, Epistemic Risk, and Overdiagnosis of Disease." *Perspectives on Science* 24 (2): 192–205.

Brigandt, Ingo. 2015. "Social Values Influence the Adequacy Conditions of Scientific Theories: Beyond Inductive Risk." *Canadian Journal of Philosophy* 45 (3): 326–56.

Brown, Matthew J. 2013. "Values in Science beyond Underdetermination and Inductive Risk." *Philosophy of Science* 80 (5): 829–39.

Brown, Matthew J., and Joyce C. Havstad. 2017. "The Disconnect Problem, Scientific Authority, and Climate Policy." *Perspectives on Science* 25 (1): 67–94.

Carrier, Martin. 2011. "Underdetermination as an Epistemological Test Tube: Expounding Hidden Values of the Scientific Community." *Synthese* 180 (2): 189–204.

Cartwright, Nancy. 2011. "A Philosopher's View of the Long Road from RCTs to Effectiveness." *Lancet* 377 (9775): 1400–1401.

Cartwright, Nancy, Jordi Cat, Lola Fleck, and Thomas Uebel. 1996. *Otto Neurath: Philosophy between Science and Politics*. New York: Cambridge University Press.

Cat, Jordi, Hasok Chang, and Nancy Cartwright. 1991. "Otto Neurath: Unification as the Way to Socialism." In *Einheit Der Wissenschaften*, ed. Jürgen Mittelstraß, 91–110. New York: de Gruyter.

Churchman, C. West. 1948. *Theory of Experimental Inference*. New York: Macmillan.

———. 1956. "Science and Decision Making." *Philosophy of Science* 23 (3): 247–49.

Cranor, Carl F. 1993. *Regulating Toxic Substances: A Philosophy of Science and the Law*. New York: Oxford University Press.

Dewey, John. 1939. "Theory of Valuation." *International Encyclopedia of Unified Science* 2 (4): 1–66.

Douglas, Heather. 2000. "Inductive Risk and Values in Science." *Philosophy of Science* 67 (4): 559–79.

———. 2009. *Science, Policy, and the Value-Free Ideal*. Pittsburgh: University of Pittsburgh Press.

———. 2016. "Values in Science." In *Oxford Handbook in the Philosophy of Science*, ed. Paul Humphreys. New York: Oxford University Press.

Elliott, Kevin. 2011. *Is a Little Pollution Good for You? Incorporating Societal Values in Environmental Research*. New York: Oxford University Press.

Elliott, Kevin, and Daniel J. McKaughan. 2014. "Nonepistemic Values and the Multiple Goals of Science." *Philosophy of Science* 81 (1): 1–21.

———. 2017. *A Tapestry of Values: An Introduction to Values in Science*. Oxford: Oxford University Press.

Giere, Ronald N. 2003. "A New Program for Philosophy of Science?" *Philosophy of Science* 70 (1): 15–26.

Goldenberg, Maya J. 2015. "How Can Feminist Theories of Evidence Assist Clinical Reasoning and Decision-Making?" *Social Epistemology* 29 (1): 3–30.

https://doi.org/10.1086/699191 Published online by Cambridge University Press

Harding, Sandra. 1991. *Whose Science? Whose Knowledge? Thinking from Women's Lives*. Ithaca, NY: Cornell University Press.
Hempel, Carl G. 1965. "Science and Human Values." In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, 81–96. New York: Free Press.
Hicks, Daniel J. 2014. "A New Direction for Science and Values." *Synthese* 191 (14): 3271–9f5.
Howard, Don. 2003. "Two Left Turns Make a Right." In *Logical Empiricism in North America*, vol. 18, ed. Gary L. Hardcastle and Alan W. Richardson, 25–93. Minneapolis: University of Minnesota Press.
———. 2006. "Lost Wanderers in the Forest of Knowledge: Some Thoughts on the Discovery-Justification Distinction." In *Revisiting Discovery and Justification*, ed. Jutta Schickore and Friedrich Steinle, 3–22. Dordrecht: Springer.
Intemann, Kristen. 2005. "Feminism, Underdetermination, and Values in Science." *Philosophy of Science* 72 (5): 1001–12.
———. 2015. "Distinguishing between Legitimate and Illegitimate Values in Climate Modeling." *European Journal for Philosophy of Science* 5 (2): 217–32.
Jeffrey, Richard C. 1956. "Valuation and Acceptance of Scientific Hypotheses." *Philosophy of Science* 23 (3): 237–46.
Jukola, Saana. 2015. "Longino's Theory of Objectivity and Commercialized Research." In *Empirical Philosophy of Science: Introducing Qualitative Methods into Philosophy of Science*, ed. Susann Wagenknecht, Nancy J. Nersessian, and Hanne Andersen, 127–43. Dordrecht: Springer.
Kourany, Janet A. 2003a. "A Philosophy of Science for the Twenty-First Century." *Philosophy of Science* 70 (1): 1–14.
———. 2003b. "A Philosophy of Science for the Twenty-First Century? Reply to Giere." *Philosophy of Science* 70 (1): 22–26.
———. 2010. *Philosophy of Science after Feminism*. New York: Oxford University Press.
Kuhn, Thomas S. 1977. "Objectivity, Value Judgment, and Theory Choice." In *The Essential Tension*, 320–39. Chicago: University of Chicago Press.
Laudan, Larry. 1984. *Science and Values: The Aims of Science and Their Role in Scientific Debate*. Berkeley: University of California Press.
Levi, Isaac. 1960. "Must the Scientist Make Value Judgments?" *Journal of Philosophy* 57 (11): 345–57.
Lloyd, Elisabeth A. 2005. *The Case of the Female Orgasm: Bias in the Science of Evolution*. Cambridge, MA: Harvard University Press.
Lloyd, Elisabeth A., and Naomi Oreskes. 2018. "Climate Change Attribution: When Is It Appropriate to Accept New Methods?" *Earth's Future* 6 (3): 311–25.
Lloyd, Elisabeth A., and Vanessa J. Schweizer. 2014. "Objectivity and a Comparison of Methodological Scenario Approaches for Climate Change Research." *Synthese* 191 (10): 2049–88.
Longino, Helen E. 1990. *Science as Social Knowledge*. Princeton, NJ: Princeton University Press.
———. 1996. "Cognitive and Non-cognitive Values in Science: Rethinking the Dichotomy." In *Feminism, Science, and the Philosophy of Science*, ed. L. H. Nelson and J. Nelson, 39–58. Dordrecht: Springer.
Mann, Michael E., Elisabeth A. Lloyd, and Naomi Oreskes. 2017. "Assessing Climate Change Impacts on Extreme Weather Events: The Case for an Alternative (Bayesian) Approach." *Climatic Change* 144 (2): 131–42.
McMullin, Ernan. 1982. "Values in Science." In *PSA 1982: Proceedings of the 1982 Biennial Meeting of the Philosophy of Science Association*, vol. 2, ed. Peter D. Asquith and Thomas Nickles, 3–28. East Lansing, MI: Philosophy of Science Association.
Melo-Martin, Inmaculada de, and Kristen Intemann. 2016. "The Risk of Using Inductive Risk to Challenge the Value-Free Ideal." *Philosophy of Science* 83 (4): 500–520.
Morrison, Monica. 2018. "Influences on Climate Modeling Methodologies." PhD diss., Indiana University.
Neurath, Otto. 1913/1983. "The Lost Wanderers of Descartes and the Auxiliary Motive (On the Psychology of Decision)." In *Philosophical Papers, 1913–1946*, ed. and trans. Robert Sonné Cohen and Marie Neurath, 1–12. Dordrecht: Springer.
———. 1928/1973. *Empiricism and Sociology*. Ed. and trans. Marie Neurath and Robert S. Cohen. Dordrecht: Springer.

https://doi.org/10.1086/699191 Published online by Cambridge University Press

Okruhlik, Kathleen. 2004. "Logical Empiricism, Feminism, and Neurath's Auxiliary Motive." *Hypatia* 19 (1): 48–72.
Osimani, Barbara, and Fiorenzo Mignini. 2015. "Causal Assessment of Pharmaceutical Treatments: Why Standards of Evidence Should Not Be the Same for Benefits and Harms." *Drug Safety* 38 (1): 1–11.
Reisch, George A. 2005. *How the Cold War Transformed Philosophy of Science: To the Icy Slopes of Logic*. Cambridge: Cambridge University Press.
Rooney, Phyllis. 1992. "On Values in Science: Is the Epistemic/Non-epistemic Distinction Useful?" In *PSA 1992: Proceedings of the 1992 Biennial Meeting of the Philosophy of Science Association*, vol. 1, ed. David L. Hull and Micky Forbes, 13–22. East Lansing, MI: Philosophy of Science Association.
Rudner, Richard. 1953. "The Scientist qua Scientist Makes Value Judgments." *Philosophy of Science* 20 (1): 1–6.
Shrader-Frechette, Kristin S., and Earl D. McCoy. 1993. *Method in Ecology: Strategies for Conservation*. Cambridge: Cambridge University Press.
Stanford, Kyle. 2016. "Underdetermination of Scientific Theory." In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Stanford, CA: Stanford University. https://plato.stanford.edu/archives/spr2016/entries/scientific-underdetermination/.
Steele, Katie. 2012. "The Scientist qua Policy Advisor Makes Value Judgments." *Philosophy of Science* 79 (5): 893–904.
Stegenga, Jacob. 2016. "Hollow Hunt for Harms." *Perspectives on Science* 24 (5): 481–504.
Uebel, Thomas E. 2005. "Political Philosophy of Science in Logical Empiricism: The Left Vienna Circle." *Studies in History and Philosophy of Science* A 36 (4): 754–73.
Wilholt, Torsten. 2009. "Bias and Values in Scientific Research." *Studies in History and Philosophy of Science* A 40 (1): 92–101.
Winsberg, Eric. 2012. "Values and Uncertainties in the Predictions of Global Climate Models." *Kennedy Institute of Ethics Journal* 22 (2): 111–37.
Zietsch, Brendan P., and Pekka Santtila. 2013. "No Direct Relationship between Human Female Orgasm Rate and Number of Offspring." *Animal Behaviour* 86 (2): 253–55.

https://doi.org/10.1086/699191 Published online by Cambridge University Press