Chapter 7
The Gap and the Error Arguments for Values in Science: Structure and Mutual Relationship

Philippe Stamenkovic

Abstract The underdetermination or 'gap' argument (GA), and the inductive risk or 'error' argument (EA), are the two main arguments used to defend the influence of non-epistemic values on scientific reasoning. However, they are often presented in a superficial or imprecise way, and their mutual relationship is not clear. This article analyzes their respective structures in detail, as well as their relationship with each other. The GA considers the logical structure of non-deductive inference (and of relative observation), and claims that (value-laden) background assumptions are needed to constitute (relative) observations and for those observations to *confirm* a hypothesis. It does not explicitly consider the *consequences* of these (value-laden) choices: rather, values are considered given and preexisting, so to speak. It is *not normative*: it claims that non-epistemic values are necessary to determine background assumptions in the sense that they are inevitable. By contrast, the EA considers the decision-theoretical problem of *accepting* or not a hypothesis, according to a (value-laden) required degree of confirmation. It is explicitly concerned with the *consequences* of these choices. It is *normative*: it claims that scientists should (in the sense of a moral obligation) take into account the non-epistemic consequences of their choices. By substituting the condition for hypothesis acceptance of the EA in the Bayesian account of the GA, one can provide a determinate limit when confirmation can be seen as sufficient to justify acceptance, thereby making the EA appear as a special case of the GA.

Keywords Gap argument · Error argument · Values in science · Underdetermination of theory by observation · Non-epistemic values

P. Stamenkovic
MTA Lendület Values and Science Research Group, ELTE Research Centre for the Humanities, Budapest, Hungary

P. Stamenkovic (✉)
University of Inland Norway, Rena, Norway
e-mail: philippe.stamenkovic@icloud.com

A. Karakas and A. T. Tuboly (eds.), *Scientific Mistakes and Mistaken Science*, Studies in History and Philosophy of Science 69, https://doi.org/10.1007/978-3-032-25925-7_7

7.1 Introduction

In the philosophical literature on values in science, one finds two main arguments to support the influence of non-epistemic values ('values' in the following) on scientific reasoning, hereafter summarized in a schematic way (one of the objectives of this article being precisely to formulate them accurately):

(1) The underdetermination argument, also known as the gap argument (GA),[1] according to which value-laden concepts and background assumptions are used by scientists to determine which data count as evidence for which hypotheses (Longino, 1990).
(2) The inductive risk argument, also known as the error argument[2] (EA), according to which a scientist has to consider the risk of being in error when deciding whether to accept or not a hypothesis, by either wrongly accepting an actually false hypothesis (type I error) or wrongly not accepting an actually true hypothesis (type II error): because of the societal consequences of this error, this decision is value-laden (Douglas, 2000, 2009).

Since these arguments are often referred to and (re)used by many authors in order to justify the influence of values on scientific reasoning, it is crucial to assess them rigorously. Indeed, they are most often superficially mentioned and taken for granted, without their logical structure being analyzed in detail. When attempts are made to do so, they are often formulated in imprecise language: for example, the GA states that background assumptions are needed to 'connect' (ChoGlueck, 2018, p. 705) or 'link' (Elliott, 2011, p. 62) evidence with theory. The same holds for their mutual relationship, which, in the rare cases where it is mentioned, is addressed superficially and/or vaguely, except by ChoGlueck (2018),[3] whose conclusion (that the EA is a special case of the GA) is nevertheless unsubstantiated (as we will see in Sect. 7.4).

Therefore, it is important to first clarify these arguments separately and then analyze their relationship. This is what this article purports to do. It does not provide an assessment of the validity of the arguments themselves—about which there is much to say (for a critique of these arguments, see Stamenkovic, 2024). In this paper, I want to elucidate these arguments and their mutual relationship, not whether they actually hold. This work is important for bringing conceptual clarity to the debate about values in science, regardless of the position defended in this debate. Accordingly, the plan of the article is the following: I will formulate the GA (Sect. 7.2) and the EA (Sect. 7.3) by returning to their original formulations, in order to be able to compare them and analyze their mutual relationship (Sect. 7.4). The EA has a simpler structure than the GA and will therefore require less analysis.

In the following and for the sake of simplicity, I will talk of theory T to designate all kinds of (sets of) theoretical (i.e. non-empirical) statements, such as theories,

[1] Expression first appearing in Intemann (2005).

[2] The first occurrence of this expression I have found is in Elliott (2011, p. 62).

[3] See Sect. 7.4. Elliott (2022, p. 19) acknowledges that this relationship deserves further scrutiny.

models, hypotheses, etc. Similarly, I will talk of observation O[4] to designate (sets of) observational (i.e. empirical) statements, which can be described as observations, data, measurements, evidence,[5] etc.

7.2 The Gap Argument

The GA is often mentioned in the values in science literature, but surprisingly rarely analyzed in detail. Most, if not all, authors just mention it and rely on it while never actually investigating it in detail. They often make use of vague, metaphoric formulations. For example, the GA enables one to 'bridge' the 'logical', 'evidential' or 'semantic' 'gap' (Longino 2004, p. 132) or 'space' (Brown, 2013, p. 834) between hypotheses and evidence. Or the GA enables to 'connect' (ChoGlueck, 2018, p. 705) or 'link' (Elliott 2011, p. 62) theory and evidence. It is used to determine the 'evidential relevance' of data (Longino, 2004, p. 43). The confusion arises from the fact that one does not know whether one starts from data and moves on to theory, or the converse.[6] Brown (2020, Ch. 2) includes 'Humean underdetermination', i.e. 'merely the uncertainty of any ampliative inference', in the GA, without explaining further. He seems to have in mind the problem of induction itself, i.e. the problem of justifying inductive inferences, in other words justifying confirmation (of theory by observation) as a rule of inference.

Basically, the GA states that logic and evidence alone are not sufficient to justify a hypothesis. Additional background assumptions are needed, (some of) which are value-laden, because epistemic values alone are insufficient to determine which background assumptions should be accepted.[7] For example, in evolution studies, there are two competing and gender-centered frameworks, 'man-the-hunter' and 'woman-the-gatherer', which can be used to interpret data (chipped stones) and explain advances in human evolution such as the development of cognitive abilities, cooperative behaviours and tool use (Longino, 1990, pp. 104–111). Each framework relies on gendered background assumptions, respectively that man hunting or

4 Here, I do not distinguish, as is sometimes done, between observations (of non-manipulated objects, such as stars) and experiments (about manipulated objects, such as particles in an accelerator), and include the latter in the former.

5 As Longino (1990) remarks, the use of the term 'evidence' can already presuppose the theory which the observation (the evidence) in question is supposed to confirm (evidence is often evidence *for something*). Therefore, this term does not appear as neutral (with respect to the theory being considered) as the other ones.

6 See e.g. the quote in footnote 21.

7 "I take the general lesson of underdetermination to be that any empirical reasoning takes place against a background of assumptions that are neither self-evident nor logically true. Such assumptions, or auxiliary hypotheses, are the vehicles by which social values can enter into scientific judgment" (Longino, 2004, p. 132).

woman gathering activities are taken to be the crucial behaviour adaptation responsible for those advances. Without these background assumptions, fossil data by itself is insufficient to decide which theory of human evolution to accept.

In fact, the GA essentially relies on two well-known issues in the philosophy of science:

- the theory-ladenness of observation;
- the underdetermination of theory by observation;

and uses these features to justify the recourse to values. These features originally appear in Duhem (1906/1981), Neurath (1913/1983),[8] and Quine (1951/1980). Longino (1990, Ch. 3) is the originator of the GA as it is used in the literature, while Intemann (2005) has analyzed it in some detail.

7.2.1 The Origins of the GA

Duhem is famous in analytic philosophy of science as the father of underdetermination. In his 1906 *La théorie physique, son objet, sa structure,* he conceptualizes four issues which are relevant here. Although Duhem only considers physical science, one can easily extend his conception to all the sciences, for example by dispensing with the symbolic formalism, or experimentation.

First, Duhem (1906/1981, pp. 221–222) presents what is now called the *theory-ladenness of observation*: "A physical experiment is the precise observation of a group of phenomena accompanied by the INTERPRETATION of these phenomena; this interpretation substitutes the concrete data really gathered with abstract and symbolic representations corresponding to them in accordance with the theories accepted by the observer" (my translation, emphasis in the original). Note that these theories are not only those presupposed by the measuring instruments, but also other theories (related to the disciplinary field investigated or to other fields) presupposed by the experimental set up. As such, theory-ladenness can be seen as a particular case of *holism* (as Longino (1990, p. 27) remarks).[9]

Elsewhere, Duhem expresses another kind of *holism*,[10] which keeps the distinction between theoretical and empirical statements: "the comparison [between theory and observation] takes necessarily place between the *whole* of theory and the *whole*

[8] Although Neurath is the only one to invoke values (although not in a clear way, see Howard (2003, 2006)), I will not spend time on his account, which I find less compelling.

[9] Later, Duhem (1906/1981, pp. 226–229) distinguishes "brute" or "practical facts" (the direct observation made in profane language) from "scientific" or "theoretical facts" (an abstract formula formulated in the symbolic language of physics). A theoretical fact can be realized experimentally by infinitely many practical facts, and conversely a practical fact can correspond to infinitely many, and logically incompatible, theoretical facts.

[10] It is misleading to label it 'confirmational holism', as Stanford (2021, sec. 2.1) does. Duhem only writes about "compar[ing]" theory and observations, in other words *testing* the theory, which can lead either to its *confirmation* or to its *disconfirmation.*

of experimental facts" (1906/1981, pp. 316–317, my translation, emphasis in the original). A single hypothesis has no observational consequence by itself (1906/1981, p. 306), and only the whole system of the physical theories used can be compared to, and tested against, the whole set of observations (1906/1981, p. 303). Without delving into the details of his conventionalism, Duhem (1906/1981, pp. 332–333) also explicitly states that logic alone and "a purely deductive method" are not sufficient to "choose" a hypothesis.

Duhem's conception of underdetermination deserves particular scrutiny (1906/1981, p. 284):

> [...] the physicist can never subject an isolated hypothesis to experimental test, but only a whole group of hypotheses; when the experiment is in disagreement with his predictions, what he learns is that at least one of the hypotheses constituting this group is unacceptable and ought to be modified; but the experiment does not designate which one should be changed.

Stanford (2021, sec. 2.1) calls this type of underdetermination "holist underdetermination", as opposed to "contrastive underdetermination". Both types of underdetermination illustrate the underdetermination of theory by evidence, and how background assumptions are needed to test a theory:

1. "Holist underdetermination" illustrates the case where an observation *disconfirms* a theory and, because hypotheses cannot be tested individually (holism) but are always conjoined to other hypotheses, one does not know which hypothesis to drop: it is therefore natural to call it *disconfirmation underdetermination*[11];
2. "Contrastive underdetermination" illustrates the case where an observation *confirms* a theory and there may be other theories equally confirmed by the same observation: let us call it more accurately *confirmation underdetermination.*[12]

Duhem spends much less time addressing confirmation underdetermination than disconfirmation underdetermination. He does so when he criticizes the "experimentum crucis" in physics (1906/1981, pt. II, ch. 6, §3). According to Duhem, crucial experiments, which are supposed to enable one to choose between two competing hypotheses, are impossible in physics. This is because in physics, contrary to the 'reductio ad absurdum' in geometry (where, if two theorems are contradictory and one is false, then the other is necessarily true), it is impossible to be certain that one has thought of all possible hypotheses to explain a phenomenon. Hence, even if, between two competing hypotheses, a set of observations experimentally disconfirms

[11] I find this expression more accurate than "holist underdetermination", which could be about confirmation as well. This expression is understandable since this variety of underdetermination is based on holism, contrary to what Stanford calls "contrastive underdetermination" (what I call *confirmation underdetermination*). Nevertheless, I find it imprecise and somewhat confusing.

[12] According to Kourany (2003a, p. 10), confirmation underdetermination, where more than one theory is confirmed by observations, happens "frequently in practice", whereas this is in fact not the case. Disconfirmation underdetermination happens much more often (Stanford, 2021). Indeed, one usually does not have enough hypotheses explaining the data, rather than having too many. Then, in the face of disconfirming evidence, one must find which hypothesis must be dropped, which is difficult because hypotheses are interconnected and cannot be easily tested in isolation.

one hypothesis, that does not mean that the other hypothesis is confirmed, because there may be a third hypothesis (or indeed many others) also explaining the same set of observations.

Quine (1951/1980) radically extends Duhem's conception of underdetermination by including: (1) statements from any science, not just physics but also formal and social science; (2) empirical statements and not just theoretical statements. He criticizes "the supposition that each statement, taken in isolation from its fellows, can admit of confirmation or infirmation at all", and claims instead that "our statements about the external world face the tribunal of sense experience not individually but only as a corporate body" (1951/1980, p. 41). Thus, "[t]he unit of empirical significance is the whole of science" (Quine, 1951/1980, p. 42). According to Quine, "it is misleading to speak of the empirical content of an individual statement" (1951/1980, p. 43). "Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system"; conversely, "no statement is immune to revision", even empirical statements or statements from logics of mathematics (1951/1980, p. 43). "Total science, mathematical and natural and human, is similarly but more extremely underdetermined by experience" (1951/1980, p. 44).

Hence, we find in Quine a radical version of disconfirmation underdetermination—but not of confirmation underdetermination.[13] In fact, Quine's holism is so extreme that it becomes difficult to even talk of disconfirmation underdetermination, or theory-ladenness of observation, because according to Quine there are simply no such things as observational (or theoretical, for that matter) statements. Because observational statements cannot be distinguished from theoretical statements, there is no confirmation problem to solve in the first place, and no theory to look for.[14]

Let us summarize the aspects of Duhem's conception which will be of interest for later, where holism can be seen as the origin of two other features:

1. *Holism*: theoretical and empirical statements are interconnected:
    (a) either across the theoretical and empirical categories (call it inter-holism): this leads to *theory-ladenness of observation*, according to which observations presuppose theories;

[13] According to Stanford (2021, sec. 3.1), Quine's disconfirmation underdetermination is mixed with his confirmation underdetermination: "After all, on Quine's view we simply revise the web of belief in response to recalcitrant experience, and so the suggestion that there are multiple possible revisions of the web available in response to any particular evidential finding just is the claim that there are in fact many different 'theories' (i.e. candidate webs of belief) that are equally well-supported by any given body of data." In a footnote, he adds: "Because the two problems are so tightly linked in Quine's epistemology, it is perhaps understandable that he gives no independent argument for taking contrastive underdetermination seriously: in what is usually cited as his most famous defence of contrastive underdetermination (though not usually distinguished from the holist variety), he simply announces, 'Surely there are alternative hypothetical substructures that would surface in the same observable ways'" (Quine, 1975, p. 313).

[14] Quine's conceptual mixture seems rather unrealistic and would probably make scientific practice impossible.

(b) or within the theoretical category (call it intra-theoretical holism): this leads to *disconfirmation underdetermination*, according to which when a phenomenon predicted by a hypothesis is not produced, one does not know whether this hypothesis or (one of) the auxiliary hypotheses is (are) disconfirmed.

2. *Confirmation underdetermination*: when a hypothesis is confirmed by some observation(s), there may be other hypotheses confirmed by the same observation(s).[15]

Note that for both disconfirmation and confirmation underdetermination, different results can be obtained if one changes the presupposed hypotheses, i.e. background assumptions (see Sect. 7.2.2):

1. A change of background assumptions can avoid disconfirming the hypothesis / theory.
2. A change of background assumptions can confirm another hypothesis / theory.

Stanford (2021, sec. 3.1) correctly advises to

> think of holist underdetermination as starting from a particular theory or body of beliefs and claiming that our revision of those beliefs in response to new evidence may be underdetermined, while contrastive underdetermination instead starts from a given body of evidence and claims that more than one theory may be well-supported by that very evidence.

Indeed, disconfirmation underdetermination starts from a *given theory* to the observation predicted by this theory; whereas confirmation underdetermination starts from a *given observation* and asks which other theories may explain this observation. Nevertheless, both types of underdetermination illustrate the *test*[16] of theory against observation—whether this test leads to the confirmation or disconfirmation of the theory—and in this sense ultimately move *from observation to theory,* which is to be justified. Disconfirmation underdetermination only starts from a provisional theory, which indeed has to be revised in the face of disconfirming observation.

7.2.2 *Contemporary Formulations of the GA*

7.2.2.1 Longino

Longino (1990) mentions three issues in her seminal book (1990, ch. 3), as well as in her other works (e.g. 2004, p. 131; 2008, pp. 69–70). She asks what determines:

[15] This is a reformulated and truncated version of Duhem's argument, suppressing its first, disconfirmation part, as we have seen. The entire argument reads: if, between two competing hypotheses, a set of observations experimentally disconfirms one hypothesis, that does not mean that the other hypothesis is confirmed, because there may be a third hypothesis (or indeed many others) also explaining the same set of observations.

[16] This aspect appears very clearly in Duhem (1906/1981, pp. 282–284).

1. "that something [i.e. a 'state of affairs' or a 'fact'] is or is taken to be evidence in the first place" (1990, p. 40)[17];
2. "whether or not someone will take some fact or alleged fact, *x*, as evidence for some hypothesis, *h*" (1990, p. 41);
3. that this fact *x* is taken as evidence for some hypothesis *h* rather than for another, "different and potentially conflicting" (1990, p. 43) hypothesis *h*'.[18]

An alternative but similar version of these questions concerns an "aspect" of a state of affairs' (p. 42) rather than the state of affairs itself. The reasoning is similar (as Longino (1990, pp. 42–43) herself acknowledges[19]) and can be analyzed in the same way, by considering an "aspect of a state of affairs" itself as a (redefined) state of affairs. Longino claims that all three questions have the same answer, namely value-laden 'background assumptions' (or, equivalently, 'background beliefs').[20]

In fact, according to Longino all three claims (that background assumptions determine 1, 2 and 3) derive from the fact that evidence and hypothesis are not formulated in the same language, in other words that the statements expressing them contain different terms. Scientific hypotheses are statements involving "such putative items as atoms, neutrinos, quarks, et cetera", whereas the evidence for them is described by statements, not about such entities, but about "cloud chambers, lines observed in spectrographic analysis, et cetera" (1990, p. 24). Indeed, if theoretical and observational statements were formulated in the same language, the underdetermination problem would amount to the problem of induction.[21] This terminological discrepancy, which

[17] This question is not really treated separately from the second one, in spite of what Longino announces. This is understandable, since, as remarked previously (footnote 5), talk of 'evidence' already presupposes a hypothesis to be confirmed. By contrast, the question of what constitutes an *observation* is addressed by Longino when she mentions "aspects of a state of affairs", which constitute the "description" of the object of study (1990, p. 42).

Note that Longino only means here the interpretation of data (which is indeed influenced by value-laden background assumptions), not its collection and characterization (which are not), as her rejection of Kuhn's (1962) or Feyerabend's (1975/1993) holism of observation makes clear (see footnote 23).

[18] Here again, the distinction between this claim and the previous one is not very clear—cut: claim 3 repeats claim 2 and adds the possibility of another hypothesis which may already be implicitly included in claim 2.

[19] "A quite different sort of example shows how different aspects of the same state of affairs can be taken as evidence for the same hypothesis, or, of course, for different hypotheses. Just as states of affairs do not stand in unique evidential relations with hypotheses, so, too, there is not a uniquely correct description for each object of description" (Longino, 1990, p. 42).

[20] Longino (1990, p. 44) also calls background beliefs "principles of inference", although strictly speaking they are statements (like observational or theoretical statements).

[21] This is what Longino (2004, p. 131) seems to have in mind when she writes: "As long as the content of theoretical statements is not represented as generalizations of data or the content of observational statements is not identified with theoretical claims, then there is a gap between hypotheses and data, and the choice of hypothesis is not fully determined by the data. Nor do hypotheses specify the data that will confirm them. Data alone are consistent with different and conflicting hypotheses and require supplementation." Indeed, Longino then distinguishes this "underdetermination problem" from "the problem of induction" (2004, p. 131). Under this "choice of hypothesis", Longino clearly has in mind confirmation underdetermination. However, she mentions Duhem as the first philosopher

Longino (2004, p. 132) calls a "semantic gap",[22] precludes "the establishing of formal relations of derivability [between theory and data] without employing additional assumptions" (Longino, 2004, p. 132), as well as "a priori specifiable statements of evidential relevance between hypothesis statements and descriptions of data" (Longino, 2008, pp. 69–70). This semantic gap includes the theory-ladenness of observation, which was already recognized by Duhem (1906/1981) (see above), whom Longino (1990) surprisingly does not quote.[23] In fact, all of Longino's three claims were already made by Duhem (see Sect. 7.2.1). Indeed, Longino's claim 1 roughly corresponds to theory-ladenness of observation and claims 2 and 3 to confirmation underdetermination (contrary to what ChoGlueck (2018, p. 716) claims),[24] although the correspondence is not clear—cut.[25] What Longino adds is the claim that the background assumptions (corresponding to Duhem's 'auxiliary hypotheses') are value-laden.[26] Claims 2 and 3 are nicely summarized by the following excerpt:

---

to raise "the problem of induction", but we saw Duhem is mostly concerned with disconfirmation underdetermination.

[22] Longino sees this semantic gap as a different type of "underdetermination" from the "problem of the existence of empirically equivalent but inconsistent theories". But the two are intimately linked, as we see hereafter.

[23] Longino (2004, p. 131) does quote him. It is not the goal of this article to compare Longino's conception to Duhem's in detail. Let me just remark that Longino (1990, pp. 26–28) rejects Kuhn's (1962) or Feyerabend's (1975/1993) relativistic senses of theory-ladenness of observation and meaning respectively, which lead to the paradoxical incommensurability of different theories. Instead, drawing on her account of background assumptions and Mary Hesse's conception of theory-ladenness of meaning, Longino (1990, pp. 53–56) claims that theory-ladenness does not imply the incommensurability of "incompatible" theories (i.e. theories giving different accounts of the same state of affairs). This is because each theory-laden set of observations provides one description (or "aspect")—among several possible—of the same state of affairs, corresponding to different background assumptions applied to this same state of affairs. Each theory therefore gives an account of a different set of evidence—and not of the same one (although these sets all relate to the same state of affairs). Therefore, there is no incompatibility. What is more, following Hesse, different theories have experiential "areas of intersection" where it is possible to compare them.

Unfortunately, Longino somewhat contradicts herself: on the one hand, she acknowledges that observations can be used as independent tests of theories as long as they are not laden with the same theory they are intended to confirm (Longino, 1990, p. 56); on the other hand, she maintains that there is an infinite regress in the background assumptions which one presupposes, whose truth cannot be evaluated (Longino, 1990, p. 52). But contrary to what Longino claims, there is nothing which prevents the background assumptions from being themselves evaluated *empirically*, as Ruphy (2006) has shown in principle, and as the history of science shows in fact (see Hacking 1983). In line with Longino's (1990, p. 56) first option just mentioned, Hacking (1983) has indeed shown how one can test new theories by using previously well-established theories (see Stamenkovic, 2022). In this way, there is no need to consider *different* sets of (theory-laden) evidence, as Longino claims: they can be empirically evaluated and accepted or rejected. If accepted, the *same* set will then be used to test competing theories.

[24] He claims that she is concerned with holist, i.e. disconfirmation, underdetermination.

[25] For example, Longino claims that (value-laden) background assumptions influence the selection of data, which is not exactly Duhem's position, who only writes that data presuppose theory.

[26] Longino more generally claims that values influence scientific practice, e.g. influence the selection of phenomena to be investigated (Longino, 1990, p. 86), but these issues fall outside the scope of the GA.

> [...] how one determines evidential relevance, why one takes some state of affairs as evidence for one hypothesis rather than for another, depends on one's beliefs, which we can call background beliefs or assumptions. Thus, a given state of affairs can be taken as evidence for the same hypothesis in light of different background beliefs, and it can be taken as evidence for quite different and even conflicting hypotheses given appropriately conflicting background beliefs. Similarly, different aspects of one state of affairs can be taken as evidence for the same hypothesis in light of differing background beliefs, and they can serve as evidence for different and even conflicting hypotheses given appropriately conflicting background beliefs. (Longino, 1990, p. 43)

To summarize Longino's position in the terminology of this article, she claims that (value-laden) background assumptions are needed:

- to constitute an observation in the first place (claim 1);
- to justify, i.e. for an observation to confirm, a theory (claim 2) including when it is competing with another one (claim 3).

Because these background assumptions are themselves value-laden,[27] values are therefore necessary to scientific reasoning. Note that Longino is not concerned with the problem of *accepting* a hypothesis (i.e. whether a hypothesis is considered true or not): rather, she is only concerned with *confirming* a hypothesis (by a given observation), as claims 1–3 make clear.

7.2.2.2 Intemann

For her part, Intemann (2005, p. 1002) distinguishes between (1) the "Duhem-Quine thesis' and (2) 'Quine's thesis of underdetermination": "Each of these theses is taken to give rise to a 'gap' that needs to be filled [by non-epistemic values] in order to be justified in accepting one theory over its competitors." With such formulation, Intemann seems to refer to confirmation underdetermination. Firstly, according to Intemann (2005, p. 1002), "The Duhem-Quine thesis is that no hypothesis, taken by itself, has any observational consequences". This leads her to her first formulation of the GA:

> G1. There is a gap between theory and observation such that auxiliary hypotheses, or background assumptions, are needed to derive testable predictions and interpret observations. In other words, background assumptions are needed to generate evidential relations between a theory and observations. (Intemann 2005, p. 1002)

[27] See e.g. (Longino, 1990, p. 86), where Longino explains that extra-scientific values can influence:

- data: "value-laden terms may be employed in the description of experimental or observational data" (claim 1, theory-ladenness of observation);
- background assumptions: "contextual [i.e. non-epistemic] values can be expressed in or motivate the background assumptions facilitating inferences" (claims 2 and 3).

Why, and how non-epistemic values are incorporated into background assumptions according to Longino, are issues left outside the scope of this article.

This formulation is somewhat unclear (especially the expression 'evidential relations'). It seems to refer to Duhem's and Quine's holisms (both inter-and intra-theoretical), but as we have seen in Sect. 7.2.1, this is only a partial account of their conception of underdetermination. Secondly:

> A second way of thinking about the gap between theory and observation arises from Quine's underdetermination thesis that there will always be multiple hypotheses (inconsistent with each other) that are consistent with all of the evidence we have at any point in time. So,
> G2. There is a gap between theory and evidence. Justification cannot simply be a logical relation between theory and evidence. There must be other characteristics of theories that make us justified in believing or rejecting a hypothesis. (Intemann 2005, pp. 1002–1003)

The first paragraph seems to concern more explicitly confirmation underdetermination. Strangely, Intemann associates this thesis with Quine, whereas we have seen that he is much more interested in disconfirmation underdetermination. What is more, G2 is in fact a different and rather vague thesis, apparently alluding to value-laden background assumptions, or perhaps to values themselves. The difference between G1 and G2 is not clear. What is more, there are some confusions in Intemann's account between (value-laden) background assumptions, value judgements, and values themselves. Intemann (2005, p. 1003) seems to confuse value-laden background assumptions with value judgements, when she writes that, according to proponents of the GA, "contextual [i.e. non-epistemic] value judgments sometimes operate as auxiliary hypotheses."

But of course, background assumptions need not be value judgements themselves, they can just be value-laden, i.e. influenced by values. For example, some values may lead to adopt certain background assumptions rather than some others, in line with Intemann's "causal interpretation" below. In fact, background assumptions need not be value-laden at all, as Intemann (2005, p. 1009) is well aware of,[28] but this issue falls outside the scope of this article. Intemann (2005, p. 1003) even seems to directly identify value-laden background assumptions with values, when she describes "contextual values" as "play[ing] the role of background beliefs in theory choice."

Intemann then distinguishes "three interpretations of how contextual values might operate to fill the gap between theory and observation", in other words three different roles which values can have in the GA. She does not relate these three interpretations to her previous G1 and G2 versions (whose usefulness in her argumentation is therefore unclear), and the interpretations themselves are not very clear neither.

1. In the "causal interpretation", "contextual values operate as *causal influences* in scientific reasoning. Contextual values may cause scientists or scientific communities to interpret data in certain ways, or to rely on certain background assumptions as opposed to others. Such values might cause scientists to give more weight

[28] For Intemann (2005, pp. 1009–1010), a way to salvage value-ladenness is to show that either the content of scientific theories, or the goal(s) pursued by science, are value-laden. But neither case depends on the existence of a "logical gap between theory and observation" (2005, p. 1011). This would therefore undermine the cogency of the GA, but again, the goal of the present article is not to evaluate the cogency of the GA, but only its structure and formulation, as well as its relationship to the EA.

to one constitutive [i.e. epistemic] value over another, or to prefer certain ways of applying or adjudicating constitutive value judgments" (2005, pp. 1004–1005). In this view, non-epistemic values "do not provide *reasons* for accepting or rejecting a theory", they "merely cause scientists to identify the negative influences of contextual values in scientific reasoning" (2005, pp. 1004–1005, original emphasis).

2. In the "tie-breakers interpretation", non-epistemic values "act [...] as 'tie-breakers' in cases where two hypotheses are equally supported by the evidence". "According to this interpretation, contextual values operate as reasons (rather than causes of reasons)[29] for taking one theory to be justified over another" (2005, p. 1007).
3. The 'normative interpretation', according to which "contextual value judgments can fill the gap by operating as background beliefs[30] in theory justification" (2005, p. 1008). "On this view, they can operate as auxiliary hypotheses in generating evidence[31] for or against a theory, and they can provide us with reasons for justifying, interpreting, applying, and adjudicating constitutive values." The difference with the causal interpretation is that here, non-epistemic values "*give us good reason* to interpret observations in a particular way, to rely on or reject a particular framework, to give more weight to some constitutive value over another, or to adopt a certain standard of evidence[32]" (2005, p. 1008, original emphasis).

Without further analyzing these interpretations which fall outside the scope of this article, one can simply summarize interpretations 1 and 3 by saying that background assumptions are *functions* of values, while safely excluding interpretation 2 as a viable option for the GA.

7.2.3 Clear Formulation

The GA, in its seminal formulation by Longino (1990), which is the version most referred to in the literature on values in science, contains the following features:

1. theory-ladenness of observation;
2. confirmation underdetermination;
3. background assumptions (used in theory-ladenness and confirmation underdetermination) are value-laden.

[29] This weird formulation apparently refers to the fact that in the tie-breaker interpretation, values directly act as a reason to accept a hypothesis (between two otherwise equally supported hypotheses, assuming such a situation is possible), whereas in the causal interpretation they act as an indirect reason influencing the choice of background assumptions or data.

[30] Again, we see the confusion between values and background assumptions.

[31] Again, a rather strange formulation.

[32] Intemann mixes here the EA with the GA. The EA is addressed in Sect. 7.3.

Thus, the GA is not an argument in the strict logical sense (i.e. an inference between premises and a conclusion), but rather a set of several claims. Note that disconfirmation underdetermination is not part of the GA. As far as I know, the GA only appears in qualitative, not quantitative (e.g. Bayesian or statistical) form in the literature, contrary to the EA. As we have seen, the GA is concerned with the *confirmation* (or *support*[33]), of theory by observation, not its *acceptance*. In the following I use the formalism presented in Lutz (2023). I use the symbol $\vdash$ for deductive inferences (entailments) and $\Vdash$ for non-deductive inferences (confirmations).

7.2.3.1 Theory-Ladenness of Observation

That observations are generally theory-laden in science is just another way of saying that they are relative rather than absolute[34]:

- absolute observation(s) (statements) are statements on which "witnesses will agree on the spot, [...] if they are conversant with the language. Their verdicts do not vary with variations in their past experience" (Quine 1975, p. 315). In other words, absolute observations are theory-independent, they are formulated thanks to sensory perceptions (e.g. "the color of this solution is red").
- by contrast, relative observations are observations relative to background assumptions: a statement O_R is an observation statement relative to background assumptions B if and only if its truth value (or its probability, in Bayesianism) follows from absolute observation statements O_A and B (Lutz 2023, p. 174). In other words, relative observations are described by statements which can be inferred from absolute observations and background assumptions (such as typically instrument readings): $O_A \wedge B \vdash O_R$.

7.2.3.2 Confirmation Underdetermination

Confirmation underdetermination is a direct consequence of the definition of hypothetico-deductive (HD) confirmation, which is the confirmation account used in the GA.[35] Assume that B are background assumptions,[36] T is a theory, O an observation. According to the HD account of confirmation, if theory T together with the background assumptions B entails observation O so that all background assumptions

[33] The latter expression makes clearer that the theory arrived at is not fully certain, in accordance with non-deductive inference.

[34] Observations of "scientific facts" (Stamenkovic, 2022) are generally relative, although they may also be absolute if they require no background assumptions.

[35] Judging from Longino (1990), where confirmation is addressed in a purely qualitative way. A generalization of the GA to a quantitative (Bayesian) account, also making use of background assumptions, is possible and does not structurally change the results presented in this subsection.

[36] Where B can stand for the conjunction of multiple separate background assumptions $B_1 \wedge B_2 \wedge ... \wedge B_n$.

are needed for the inference, then O confirms the conjunction of T and B. More precisely (adapted from Lutz 2023, p. 98):

If

- $B \nvdash \neg T$ (i.e. T, B are consistent);
- $B \nvdash O$ (i.e. B are plausible independently of O);
- $T \wedge B \vdash O$ (where all of T, B are needed for the inference);

Then

$$O \Vdash T \wedge B.$$

In other words, if a theory is compatible with the background assumptions and entails an observation statement that is not already entailed by the background assumptions, then that observation statement confirms (supports) the conjunction of the theory and the background assumptions. By using the consequence condition of confirmation[37] (thanks to which one can confirm a theory by confirming a more specific theory) (Hempel 1945a, 1945b), we can further conclude that the observation statement O confirms the theory T.

Confirmation underdetermination becomes clear with this formalization: one may come up with a different theory $T^{'}$ (and possibly different background assumptions $B^{'}$), which entail(s) O: $T^{'} \wedge B^{'} \vdash O$ (where all of $T^{'}$, $B^{'}$ are needed for the inference). Then (given that $T^{'}$, $B^{'}$ are consistent and $B^{'}$ are plausible independently of O) O confirms $T^{'} \wedge B^{'}$, and further confirms $T^{'}$ by the consequent condition. Thus, we are unable to prefer T over $T^{'}$ if both are confirmed by O. The choice between different such theories (equally well confirmed) constitutes the value-laden aspect of the HD account of confirmation.[38]

We have just seen that the HD account of confirmation does not distinguish between two theories which entail the same observational statements. Similarly,[39] the Bayesian account of confirmation does not distinguish between two theories which assign the same probabilities to all observational statements. In Bayesian formalism, the degree of confirmation of a theory is given by its probability, and an observation O supports (or confirms) a theory T with background assumptions B if and only if (adapted from Lutz, 2023, p. 124):

$$P(T \wedge B|O) > P(T \wedge B) \tag{7.1}$$

---

[37] Which can be formulated thus (adapted from Lutz, 2023, p. 97): let T be a theory, B background assumptions and O observations. If $O \Vdash T \wedge B$ and $T \wedge B \vdash U$, then $O \Vdash U$, where U is a statement (or theory). In particular, it is obvious that $T \wedge B \vdash T$, hence $O \Vdash T$.

[38] Note that disconfirmation underdetermination also follows from the HD account of confirmation. Indeed, if O is false, $T \wedge B$ is false: $\neg O \vdash \neg(T \wedge B)$, which is equivalent to $\neg O \vdash \neg T \vee \neg B_1 \vee \ldots \vee \neg B_n$. In other words, if an observation disconfirms a theory, we do not know whether it is the theory or one of the background assumptions which is false.

[39] Although Longino does not consider the Bayesian approach, I give it here for information. Indeed, it is always possible to express an HD-account as a Bayesian account, since Bayesian confirmation entails HD confirmation (see Lutz, 2023, pp. 141–143).

7.2.3.3 Value-Ladenness of Background Assumptions

So far, Longino's GA does not add anything new. Its distinctive feature is that it considers background assumptions as *functions* of values v (it may be that values make us choose one background assumption over another, or that the content of the background assumption is changed, etc.). However, apart from her example in evolutionary biology, Longino does not explain how this function is supposed to work in general, and precisely. This dependence can be taken into account by simply writing $B(v)$ in the previous formalism.

To sum up, a theory is underdetermined by the observations that confirm it. Different theories are equally well confirmed as long as they entail the same observations. We cannot, on the mere basis of observations which confirm two different theories (and logic), choose between these theories. But if we add values, we can.

Finally, it is worth noting that the GA (whether in Longino's or in Intemann's formulation) is *not* normative: value-laden background assumptions are necessary in the sense that they are inevitable (that it is impossible to do science without them), not in the sense that they *should* be used (in a moral sense).

7.3 The Error Argument

The EA originally appeared in Churchman (1948), was clearly formulated by Rudner (1953), and was especially developed by Douglas (2000, 2009, 2017). The main difference with the GA is that the EA considers a *decision-theoretical* framework where one *accepts or not* a hypothesis. The scientist is then morally obliged to consider the non-epistemic consequences, and hence associated values, of this decision.

7.3.1 The Origins

Such a decision-theoretical approach is evident in Churchman's (1948) "theory of pragmatic inference". In contrast to the theory of statistical inference, in this theory

> [...] the scientific method is conceived as an activity designed to choose the most efficient means for one end, or a set of consistent ends. The theory of pragmatic inference depends on establishing a functional relationship between the relative efficiency of a course of action and the probability of choosing the action. (Churchman, 1948, p. 267)

According to this theory "every scientific hypothesis is considered to be a possible course of action for accomplishing a certain end, or set of ends" (1948, p. 259). Churchman also insists on all the decisions required in scientific inquiry besides hypothesis acceptance, such as in data gathering, making assumptions, etc. (1948,

p. 259): decisions which will later be considered by proponents of values (such as Douglas, 2000), and conceived as value-laden, in all phases of scientific research.

Levi (1961, p. 614) provides an interesting classification of "behavioralist" viewpoints on induction, which extend the domain of

> inductive behavior (where the problem is to formulate criteria for deciding what to do when the decision maker is uncertain about the state of affairs under which he is operating) into the domain of inductive inference (where the problem is to formulate criteria for accepting and rejecting hypotheses when the available evidence does not entail the truth or falsity of any of the hypotheses being considered).

More precisely, according to the behavioralist viewpoint, "if scientists qua scientists accept hypotheses at all they do so in a sense that equates accepting a hypothesis *H* with acting on the basis of *H* with respect to a practical objective *O*" (1961, p. 615). Within this behavioralist strand, Levi (1961, p. 615) distinguishes "the decision-maker conception of the scientist, according to which the scientist does accept and reject hypotheses, but only in a behavioral sense" (where he puts Churchman and Rudner), from "the guidance-counselor conception, according to which the scientist neither accepts nor rejects hypotheses in any sense, behavioral or nonbehavioral, but only assigns degrees of confirmation to hypotheses which may be of use to the practical decision maker" (to which Carnap, Hempel and Jeffrey belong). The EA clearly fits into the decision-making conception.

In what has become a seminal article, Rudner explains what it means that "the scientist as scientist accepts or rejects hypotheses":

> But if this is so then clearly the scientist as scientist does make value judgments. For, since no scientific hypothesis is ever completely verified, in accepting a hypothesis the scientist must make the decision that the evidence is *sufficiently* strong or that the probability is *sufficiently* high to warrant the acceptance of the hypothesis. Obviously our decision regarding the evidence and respecting how strong is strong enough, is going to be a function of the *importance*, in the typically ethical sense, of making a mistake in accepting or rejecting the hypothesis. [...] *How sure we need to be before we accept a hypothesis will depend on how serious a mistake would be.*
>
> [...] *In general then, before we can accept any hypothesis, the value decision must be made in the light of the seriousness of a mistake, that the probability is high enough or that, the evidence is strong enough, to warrant its acceptance.* (Rudner, 1953, pp. 2–3, original emphasis)

This focus on *error* in hypothesis acceptance has led philosophers to dub this argument the inductive risk argument, or indeed the error argument. Such an understanding restricted to taking into account non-epistemic consequences and values only in case of *error* in hypothesis acceptance is the one of Douglas (2000, p. 564), who revived this argument. However, the *correct* outcomes should also be taken into account and valued, as we shall see in Sect. 7.3.3. Therefore, the appellations "error argument", and even "inductive risk argument", are somewhat misleading. Such a larger understanding is the one Hempel (1965, p. 92) proposes, who talks of the "values" and "disvalues" associated with the intended and unintended outcomes of decision-making in hypothesis acceptance:

> As was noted earlier, the formulation of "adequate" decision rules requires, in any case, the antecedent specification of valuations that can then serve as standards of adequacy. The requisite valuations, as will be recalled, concern the different possible outcomes of the choices which the decision rules are to govern. Now, when a scientific rule of acceptance is applied to a specified hypothesis on the basis of a given body of evidence, the possible "outcomes" of the resulting decision may be divided into four major types:(1) the hypothesis is accepted (as presumably true) in accordance with the rule and is in fact true; (2) the hypothesis is rejected (as presumably false) in accordance with the rule and is in fact false; (3) the hypothesis is accepted in accordance with the rule, but is in fact false; (4) the hypothesis is rejected in accordance with the rule, but is in fact true. The former two cases are what science aims to achieve; the possibility of the latter two represents the inductive risk that any acceptance rule must involve. And the problem of formulating adequate rules of acceptance and rejection has no clear meaning unless standards of adequacy have been provided by assigning definite values or disvalues to those different possible "outcomes" of acceptance or rejection. It is in this sense that the method of establishing scientific hypotheses "presupposes" valuation: the justification of the rules of acceptance and rejection requires reference to value judgments. (Hempel, 1965, p. 93)

7.3.2 Douglas

Drawing on a case study of dioxin toxicity, Douglas (2000) essentially endorses Rudner's stance, which she extends not only to the acceptance/rejection stage of scientific inquiry, but also to the other stages[40]:

- choice of methodology: the statistical design of the study (the choice of the level of statistical significance), balancing the risk of false positives (leading to over-regulation and detrimental economic consequences) vs false negatives (leading to underregulation and detrimental public health consequences);
- "collection and characterization of data": in Douglas's rat liver tumors example, the judgement whether a tissue sample has a cancerous lesion or not depends on the inductive risk i.e. the consequences of potential errors; and
- "interpretation of data": in Douglas's example, whether one chooses a threshold or linear extrapolation model for dioxins' carcinogenic effects.

However, the distinction between "characterization" and "interpretation" of the data is not clear, since characterization already presupposes interpretation in Douglas's example. Both can in fact be seen as examples of hypothesis acceptance, the hypothesis being that a rat liver is tumorous, or that dioxins' carcinogenic effects follow a certain model. Douglas (2000, p. 565) also distinguishes the "interpretation of results" (2000, p. 565) from the "acceptance of hypotheses" (2000, p. 563), but this seems in fact to be one and the same thing. In her example, the interpretation of data amounts to accepting or not the hypothesis that there is a threshold in dioxins' carcinogenic effects, and ultimately that dioxins cause liver cancer in rats. Although Douglas (2000, sec. 4) considers the choice of methodology separately, it also relates in fact to hypothesis acceptance, since the level of statistical significance is used to

[40] As remarked, such a view was already sketched by Churchman (1956).

reject or not the null hypothesis. In the end, all these stages amount to hypothesis acceptance.

Background assumptions only intervene in what Douglas calls the interpretation of data. Douglas claims that they influence which interpretation one comes up with, and that the choice of these background assumptions itself depends on the risks associated with this choice:

> The background assumptions that lead to the threshold position include that the toxicity seen in the livers is a likely cause of the cancers and that such cancer promotion is likely to be a threshold phenomenon, making it more plausible that an apparent thresh- old in the data is an actual threshold. The background assumptions that lead to the opposing viewpoint include that the statistical sensitivity of the studies is not sufficient to detect a threshold and that the link between toxicity and cancer promotion is correlation but not necessarily causation. (Douglas, 2000, p. 576)

Douglas (2009, p. 96) also distinguishes two roles which values can have: a "direct role" when values "act as reasons in themselves to accept a claim"; and an "indirect role" when they help determine how much evidence to demand. Only the latter role is of course vindicated and considered here as part of the EA, which in general can be formulated as a hypothesis acceptance decision-making problem.

7.3.3 Clear Formulation

This decision-theoretical problem can be formulated in two main ways, as Levi (1962) already suggested: Bayesian statistics or null hypothesis significance testing. Both enable to quantify the level of evidence required to accept a hypothesis, a quantification which in general is presupposed by this problem. In null hypothesis significance testing, the choice of the level of statistical significance is value-laden and illustrates "the degree of caution exercised in a search for the truth" (Levi, 1962, p. 63). It is the maximum probability of committing a type I error (accepting a false proposition as true). It has to be traded-off with the desire to avoid type II errors (not accepting a true proposition). What is more, even the choice of the null hypothesis may be value-laden, since in principle it may be switched with the hypothesis, even if in practice scientists choose the null hypothesis so as to minimize type I errors.[41]

In the following, I only consider Bayesian statistics. According to the EA, social values determine how high the probability of a theoretical statement (hypothesis) T, in other words its degree of confirmation, must be relative to a body of observations O[42] in order to accept it: $P(T \vee O) > P_0(v)$ where $P_0(v)$ is a probability depending

[41] Douglas (2000) does not mention this influence of values on methodological choices.

[42] According to Douglas, the observations themselves depend on values: one might therefore want to write O = O(v). However, we have seen that this issue itself reduces to the one of hypothesis acceptance.

on values *v*. Using Bayesian decision-making based on expected utility theory, one can write, adapting from Levi (1962, sec. II)[43]:

	T is true	T is false
Accept T	u_{11}	u_{12}
Do not accept T	u_{21}	u_{22}

where u_{ij} are the utilities associated with action i (1: accept; 2: do not accept) and case j (1: T is true; 2: T is false). As mentioned previously, the correct outcomes also have to be considered. Then, using again the previous notation with theory T, observations O and background assumptions B:

- if

$$P(T \wedge B|O) > \frac{1}{1 + \frac{u_{11} - u_{21}}{u_{22} - u_{12}}} \tag{7.2}$$

  the scientist should accept T;
- if $P(T \wedge B|O) < \frac{1}{1+\frac{u_{11}-u_{21}}{u_{22}-u_{12}}}$ the scientist should not accept T.[44]

According to this enlarged version of the EA, the utilities u_{ij} are function of non-epistemic values v_{ij} associated with each case: $u_{ij} = u_{ij}(v_{ij})$ (note that they are not values themselves). They have an influence on the level of proof required for a hypothesis, i.e. its degree of confirmation, in order to accept it.

Note that contrary to the GA, which states the inevitability of non-epistemic values, the EA states the moral obligation to have recourse to such values. Therefore, only the latter is normative: scientists *should* take into account the consequences of their decisions—but they may happen not to do so (again, contrary to the GA according to which it is impossible not to use value-laden background assumptions).

7.4 Relationship Between the Two Arguments

There are few references on the relationship between the EA and the GA. According to Elliott (2011, p. 70), all three "principles"[45] which he sees as justifying the influence of non-epistemic values on science "are visible in the application of the error

[43] Contrary to Levi, I do not consider the option of suspending judgement: either one accepts a hypothesis, or one does not.

[44] If PTB|O = 11 + u11-u21u22-u12, both expected utilities associated with accepting and not accepting T are equal, and the scientist can do either.

[45] Namely: the ethical responsibility of scientists to consider the societal consequences of their work; the fact that they often face situations of uncertainty and have to decide which standards of proof to require before making a decision to accept or not a claim; and the fact that it is impractical or harmful for scientists to defer this decision to extra-scientific decision-makers (e.g. politicians) (2011, p. 55).

argument as well as the gap argument"—a rather vague formulation, which apparently means that these principles are common premises of the two arguments. Because both arguments concern "inferential gaps" in scientific reasoning, Betz (2013, p. 208) finds it "not clear [...] whether we have two distinct (albeit closely related) arguments at all", and treats them as the same "methodological" argument, according to which "scientists have to make methodological decisions which require them to rely on non-epistemic value judgments" (2013, p. 209). Biddle (2013, p. 125) views the EA as a "special case" of the GA, although he does not justify his position. According to Brown (2013, p. 834), both arguments "share a common premise" inasmuch as they "begin from a situation where the evidence is fixed and take values to play a role in the space that is left over" (what Brown criticizes is that for both, evidence has priority over values). Again, we see imprecise expressions such as the "space that is left over". Brown (2020, ch. 2) further sees both arguments as two "instances"[46] of his "contingency argument", according to which scientists have to make value judgements to take decisions in all the "contingent moments" of scientific inquiry where different "reasonable alternative[s]" or "unforced choices" are available. But he does not specify what exactly is the relation between these two instances, and his own contingency argument is too general and vague to shed light on this issue.

ChoGlueck's (2018, p. 707) account is, to my knowledge, the only attempt to investigate into detail the relationship between the GA and the EA. ChoGlueck (2018, p. 716) claims that the EA can be seen as a special case of the GA: it is "nested within the gap because the error is a limited case of the gap with narrower features" (2018, p. 704). ChoGlueck (2018, p. 704) essentially bases his analysis on the views of Neurath's (1913/1983), Longino (1990) and Kourany (2003a, 2003b, 2010), which he tries to decompose into four coherent "features" for each argument (G1 to G4 and E1 to E4), although the views of these authors are not really the same, and extend the GA so much so that it becomes too general and vague.[47] Therefore, in my account of ChoGlueck's argumentation I only refer to his use of Longino's position.

Although I more or less agree with ChoGlueck's conclusion, I do not agree with the way he reaches it. His account is also vague, makes use of imprecise formulations (typically using the same kind of expressions as those mentioned previously[48]) and is sometimes wrong. According to ChoGlueck,

> [...] 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

---

[46] More exactly, he uses this term for the underdetermination argument (which, according to him, has several different forms), and talks of "an elaboration of a specific form of the contingency argument" for the "error argument" version of the inductive risk argument, and of "a version of the contingency argument" for the "pragmatic argument" version of the inductive risk argument.

[47] ChoGlueck (2018) considers Neurath's (1913/1983, p. 4) "auxiliary motives" and Kourany's "goals", "values" and "ideals" to also fall under the GA.

[48] E.g. "a space for societal values to improve scientific knowledge".

> 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. (ChoGlueck, 2018, p. 711)

More precisely, the four "features" of the gap argument are (2018, sec. 2):

- G1: "Empirical Uncertainty", which in fact refers to (1) the theory-ladenness of observation[49] and (2) confirmation underdetermination[50];
- G2: "A Bridge Feature" between evidence and theory, i.e. background assumptions in Longino's case;
- G3: "Societal Stakes of Scientific Knowledge": background assumptions are based on non-epistemic values;
- G4: "Objectivity through Intersubjectivity": mutual criticism of background assumptions in the scientific community.

ChoGlueck summarizes the EA as follows:

> 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 necessary role in setting standards of evidence when social consequences are possible. (ChoGlueck, 2018, p. 714)

The four features of the error argument are:

- E1: "Evidential Uncertainty": contrary to what one may think, ChoGlueck (2018, pp. 712–713) does not mean uncertainty related to observations, but to scientific inference itself,[51] i.e. "the problem of induction"[52];
- E2: "Standards of Evidence" refer to the "quality and quantity of evidence" required to accept a hypothesis;
- E3: "Social Consequences of Scientific Error" refers to errors in hypothesis acceptance which have a social consequence;
- E4: "Responsibility to Consider Social Consequences" only adds the responsibility to consider E3.

[49] Which ChoGlueck (2018, p. 708) explicitly mentions.

[50] ChoGlueck quotes Longino's (1990) "logical gap" between theory and evidence (1990, pp. 43, 52).

[51] He mentions for example Hempel's "inductive risk" (1965, p. 92), and Douglas's "inductive gap" (2009, p. 96).

[52] ChoGlueck (2018, p. 713) also includes in this feature the question 'how much evidence is enough to make a general claim from knowledge of particulars?', thereby blurring the distinction between this feature and the next one (E2).

Now according to ChoGlueck (2018, p. 715) the EA can be seen "as a special, limited case" of the GA, or "as a concrete example (or set of examples)" (2018, p. 718) of the latter, because:

1. "[E]vidential 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" (2018, p. 716). However, this interpretation is not fully correct. First, the EA, like the GA, is also concerned with what counts as evidence as we have seen with Douglas. The EA adds the—rather distinct—issue of how much evidence is enough to accept a claim, which is a decision-theoretical problem, whereas the GA only states (confirmation) underdetermination. By contrast, the problem of induction consists in *justifying* confirmation as a rule of inference. For ChoGlueck (2018, p. 716), E1 illustrates what he calls "enumerative underdetermination", "which focuses on the threshold at which one can make an inference from data to theory" and which is according to him a subcategory of Stanford's "contrastive underdetermination", i.e. confirmation underdetermination in the terminology of this article. He puts Kourany and Neurath[53] in this category, but according to him Longino only illustrates "holist", i.e. disconfirmation, underdetermination, where "theoretic assumptions influence what observations even count as evidence". We have seen, however, that this interpretation is not correct: rather, this point amounts in fact to theory-ladenness of observation, and in addition Longino's position illustrates confirmation underdetermination.
2. "[S]tandards of evidence" (e.g. statistical significance levels) used in E2 are "a paradigmatic example of a bridge features G2, which connects observations to theory as supportive evidence" (2018, p. 716). However, and as ChoGlueck himself acknowledges, they are used in a "decision rule" about accepting or not a hypothesis, whereas the GA is only about hypothesis confirmation or support, which is a different issue and does not necessarily imply acceptance. As we have seen, the GA is not only about "evidential relevance" (2018, p. 716)—a rather vague expression, which has to do with theory-ladenness and holism. Conversely, the EA is not only about "evidential sufficiency", as ChoGlueck (2018, p. 716) seems to imply, but also about "evidential relevance", as we have seen with Douglas (2000) regarding data collection and characterization. What is more, saying that the issue of "evidential relevance" is "nested" within the issue of "evidential sufficiency", as ChoGlueck (2018, p. 716) writes, is also an unclear expression. Rather, as he also correctly writes, the former is (chrono)logically "prior" to the latter. But that does not mean that the latter can be considered as a "special, limited case" of the former.
3. I agree that "the claim that ethically relevant consequences may follow from scientific errors (E3) is one case of how scientific knowledge production has societal stakes (G3)" (2018, p. 717).

[53] Neurath also illustrates "holist", i.e. disconfirmation underdetermination according to ChoGlueck.

4. Finally, while "[t]he final features of each argument (G4 and E4) share the same neutral notion of individual bias as ineliminable and enabling [sic] by rejecting the detached sense of objectivity", and while "both approaches use societal values as grounds for (indirect) arbitration of evidence claims" (2018, pp. 717–718), that does not mean that they stand in a "nested" relation, but rather share common assumptions about scientific objectivity. What is more, G4 assumes the inevitability of non-epistemic values, while E4 assumes their moral necessity.[54] G4 claims that objectivity is achieved through intersubjectivity and mutual criticism, while E4 only states that (individual)[55] scientists have the responsibility to consider the social consequences of their errors. ChoGlueck (2018, p. 719) also claims that the EA is based on a "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 test", while the GA uses "an inductive/abductive model of confirmation, whereby one looks at a variety of evidence and assesses validity based on coherence and robustness". While the latter characterisation of the GA is rather vague, we have seen that it is based on confirmation underdetermination and can also be expressed according to HD confirmation. Therefore, ChoGlueck's (2018, p. 719) claim that "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" is irrelevant (not to speak about its correctness).

Thus, ChoGlueck (2018)'s conclusion that the EA is a "special", "limited" case of the GA, and an "example" of it, is unsubstantiated.

7.5 Conclusion

The GA and the EA both rely on underdetermination of theory by evidence, i.e. the fact that observation and logic alone do not enable to infer which theory is correct, and claim that values are needed for doing so. The two arguments therefore both have to do with the problem of induction and share a common premise. They also both make use of (value-laden) background assumptions, and both consider the choice of methodology to be value-laden as well (Douglas, 2000, sec. 4; cf. Longino, 1990, pp. 83–85). Thus, the GA and the EA have much in common: they both claim that the choice of methodology, data and hypothesis are value-laden.

What differentiates them is that:

- the GA:

[54] G4 even assumes epistemic improvement through non-epistemic values, which is not the case of E4.

[55] ChoGlueck (2018, p. 716) claims that the EA also has a collective dimension of openness and mutual criticism between scientists, but this depends on the authors and is not the case of Douglas (2000).

Table 7.1 Comparison of the GA and the EA

	Object	Formal framework	Influence of values	Modal status	Consequences
GA	Hypothesis confirmation	HD (or Bayesian) confirmation	Background assumptions	Necessary (inevitable)	Not taken into account
EA	Hypothesis acceptance	Bayesian decision-theory	Utilities	Normative (and descriptive)	Taken into account

- considers the logical structure of non-deductive inference (and of relative observation), and claims that (value-laden) background assumptions are needed to constitute (relative) observations and for those observations to *confirm* (or *support*) a hypothesis;
- does not explicitly consider the *consequences* of these (value-laden) choices (whether they are correct or not): rather, values are considered given and preexisting, so to speak;
- is *not normative*: it claims that non-epistemic values are necessary to determine background assumptions in the sense that they are inevitable.

• whereas the EA:

- considers the decision-theoretical problem of *accepting* or not a hypothesis, according to a (value-laden) required degree of confirmation (probability level);
- is explicitly concerned with the *consequences* of these choices, according to their correctness;
- is *normative*: it claims that scientists should (in the sense of a moral obligation) take into account the non-epistemic consequences of their choices, and corresponding values.[56]

Table 7.1 summarizes these findings:

Finally, it is possible to establish a link between both arguments. The condition (Eq. (7.2)) stating when a scientist should accept T can be used in the Bayesian account of confirmation (Eq. (7.1)) to provide a determinate limit when confirmation can be seen as sufficient to justify acceptance. In this way, one may consider the EA as a special case of the GA. This would lead to the same conclusion as ChoGlueck (2018), but not for the same reasons.

By clarifying the GA, the EA and their mutual relationship, this article will hopefully contribute to a clearer debate about the influence of values in scientific reasoning, by showing their distinct influence on hypothesis confirmation and acceptance, respectively, and their associated conceptual frameworks.

[56] According to some authors, including Douglas (2017, 83–84), the EA can also have a descriptive component considering non-epistemic values to be inevitable.

Acknowledgments My heartfelt thanks to Adam Tamas Tuboly, whose support and patience were decisive in my achievement of this article, and who kindly took care of all the editing. Many thanks to Alexandra Karakas for her detailed and helpful comments.

References

Betz, G. (2013). In defence of the value free ideal. *European Journal for Philosophy of Science, 3*(2), 207–220.
Biddle, J. (2013). State of the field: Transient Underdetermination and values in science. *Studies in History and Philosophy of Science Part A, 44*(1), 124–133.
Brown, M. J. (2013). Values in science beyond Underdetermination and inductive risk. *Philosophy of Science, 80*(5), 829–839.
Brown, M. J. (2020). *Science and moral imagination: A new ideal for values in science*. University of Pittsburgh Press.
ChoGlueck, C. (2018). The error is in the gap: Synthesizing accounts for societal values in science. *Philosophy of Science, 85*(4), 704–725.
Churchman, C. W. (1948). Statistics, pragmatics, induction. *Philosophy of Science, 15*(3), 249–268.
Churchman, C. W. (1956). Science and decision making. *Philosophy of Science, 23*(3), 247–249.
Douglas, H. (2000). Inductive risk and values in science. *Philosophy of Science, 67*(4), 559–579.
Douglas, H. (2009). *Science, policy, and the value-free ideal*. University of Pittsburgh Press.
Douglas, H. (2017). Why inductive risk requires values in science. In K. C. Elliott & D. Steel (Eds.), *Current controversies in values and science* (pp. 81–93). Routledge.
Duhem, P. (1906/1981). *La Théorie Physique: Son Objet, Sa Structure*. Vrin.
Elliott, K. C. (2011). *Is a little pollution good for you?: Incorporating societal values in environmental research*. Oxford University Press.
Elliott, K. C. (2022). *Values in science*. Cambridge University Press.
Feyerabend, P. (1975/1993). *Against method*. Verso.
Hacking, I. (1983). *Representing and intervening*. Cambridge University Press.
Hempel, C. G. (1945a). Studies in the logic of confirmation (I). *Mind, 54*(213), 1–26.
Hempel, C. G. (1945b). Studies in the logic of confirmation (II.). *Mind, 54*(214), 97–121.
Hempel, C. G. (1965). Science and human values. In *Aspects of scientific explanation; and other essays in the philosophy of science* (pp. 81–96). The Free Press.
Howard, D. (2003). Two left turns make a right: On the curious political career of north American philosophy of science at midcentury. In G. L. Hardcastle & A. W. Richardson (Eds.), *Logical empiricism in North America* (pp. 25–93). University of Minnesota Press.
Howard, D. (2006). Lost Wanderers in the Forest of knowledge: Some thoughts on the discovery-justification distinction. In J. Schickore & F. Steinle (Eds.), *Revisiting discovery and justification: Historical and philosophical perspectives on the context distinction* (pp. 3–22). Springer.
Intemann, K. (2005). Feminism, underdetermination, and values in science. *Philosophy of Science, 72*(5), 1001–1012.
Kourany, J. A. (2003a). A philosophy of science for the twenty-first century. *Philosophy of Science, 70*(1), 1–14.
Kourany, J. A. (2003b). A philosophy of science for the twenty-first century? Reply to Giere. *Philosophy of Science, 70*(1), 22–26.
Kourany, J. A. (2010). *Philosophy of science after feminism*. Oxford University Press.
Kuhn, T. (1962). *The structure of scientific revolutions*. The University of Chicago Press.
Levi, I. (1961). Decision theory and confirmation. *The Journal of Philosophy, 58*(21), 614–625.
Levi, I. (1962). On the seriousness of mistakes. *Philosophy of Science, 29*(1), 47–65.
Longino, H. E. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton University Press.

Longino, H. E. (2004). How values can be good for science. In P. Machamer & G. Wolters (Eds.), *Science, values, and objectivity* (pp. 127–142). University of Pittsburgh Press.

Longino, H. E. (2008). Values, heuristics, and the politics of knowledge. In M. Carrier, D. Howard, & J. Kourany (Eds.), *The challenge of the social and the pressure of practice: Science and values revisited* (pp. 68–86). University of Pittsburgh Press.

Lutz, S. (2023). *Philosophy for scientists and engineers*, unpublished book typescript.

Neurath, O. (1913/1983). The lost Wanderers of Descartes and the auxiliary motive. In R. S. Cohen & M. Neurath (Eds.), *Otto Neurath: Philosophical papers 1913–1946* (pp. 1–12). Springer.

Quine, W. V. O. (1951/1980). *Two dogmas of empiricism, in from a logical point of view: 9 Logico-philosophical essays* (2nd ed., pp. 20–46). Harvard University Press.

Quine, W. V. O. (1975). On empirically equivalent Systems of the World. *Erkenntnis, 9*, 313–328.

Rudner, R. (1953). The scientist qua scientist makes value judgments. *Philosophy of Science, 20*(1), 1–6.

Ruphy, S. (2006). 'Empiricism all the way down': A defense of the value-neutrality of science in response to Helen Longino's contextual empiricism. *Perspectives on Science, 14*(2), 189–214.

Stamenkovic, P. (2022). Facts and objectivity in science. *Interdisciplinary Science Reviews, 48*(2), 277–298.

Stamenkovic, P. (2024). Straightening the 'value-laden turn': Minimising the influence of values in science. *Synthese, 203*(20).

Stanford, K. (2021). Underdetermination of scientific theory. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Plato. https://plato.stanford.edu/archives/win2021/entries/scientific-underdetermination/

Philippe Stamenkovic is currently temporary associate professor at the University of Inland Norway. He previously worked at the University Paris Diderot, Aix-Marseille University, Ben-Gurion University of the Negev, and Uppsala University (and before that as an aeronautical engineer). He is interested in the relationships between science and society, especially how social values influence scientific practice, and conversely how scientific knowledge is used in society. He is also interested in research ethics.