# The Epistemic Risk in Representation

Stephanie Harvard, Eric Winsberg


Kennedy Institute of Ethics Journal, Volume 32, Number 1, March 2022, pp. 1-31 (Article)

Published by Johns Hopkins University Press
*DOI:* *https://doi.org/10.1353/ken.2022.0001*

*Stephanie Harvard and Eric Winsberg*

# The Epistemic Risk in Representation

ABSTRACT. Both the distinction between the 'internal' and 'external' phases of science and the concept of 'inductive risk' are core constructs in the values in science literature. However, both constructs have shortcomings, which, we argue, have concealed the unique significance of values in scientific representation. We defend three closely-related proposals to rectify the problem: i) to draw a conceptual distinction between endorsing a 'fact' and making a decision about representation; ii) to employ a conception of inductive risk that aligns with this distinction, not one between internal/external phases in science; iii) to conceptualize 'representational risk' as a unique epistemic risk, no less significant than inductive risk. We outline the implications of each proposal for current debates in the values in science literature.

## 1. INTRODUCTION

It is widely agreed in the 'values in science' literature that not all ways in which values play a role in science are equally epistemologically significant. In particular, many contributors attach special significance to the role of values in 'internal' (not 'external') phases of science,[1] such as in managing inductive risk (Douglas 2000; 2009; Elliott and Richards 2017). Both the distinction between the internal/external phases of science and the concept of inductive risk, then, are core constructs in the values in science literature: both help philosophers interpret the epistemological significance of values in science. At the same time, both constructs have shortcomings: several contributors have argued that the internal/external distinction is insignificant (Bueter 2015; Elliott and McKaughan 2009; Winsberg 2018, Ch.9) and the definition of inductive risk is debated (Biddle and Kukla 2017; Powers 2017; Elliott and Richards 2017b). In this paper, we consider these shortcomings, and argue that they have prevented philosophers from fully appreciating and understanding the unique significance of values in *scientific representation*. Moreover, we argue for

*Kennedy Institute of Ethics Journal* Vol. 32, No. 1, 1–31 © 2022 by Johns Hopkins University Press

three closely related proposals to help recognize this significance: i) to draw a conceptual distinction between endorsing a 'fact'[2] and making a decision about representation; ii) to employ a conception of inductive risk that aligns with this distinction, not one between internal/external phases in science; iii) to conceptualize 'representational risk' as a unique epistemic risk, no less significant than inductive risk. We show that these three proposals generate a variety of benefits.

We build our argument on the 'models in science' literature, in which it is firmly established that models are not true or false, but rather adequate or inadequate for purpose (Alexandrova 2010; Bokulich and Parker 2020; Frigg and Hartmann 2020; Parker 2010; 2020; van Fraassen 1980; Winsberg 2018; 2010). The unique way that this invites values to enter into science has yet to be fully articulated.[3] In what follows, we show that adequacy for purpose, and its attendant value-ladenness, applies broadly to *representational decisions* in science. In our terminology, representational decisions include decisions about 'what to represent' (i.e., decisions about what entities to include in and exclude from a representation) and 'how to represent' (i.e., decisions about entities already chosen for inclusion in the representation), whether in a model or another representational device. For example, we take the choice of comparator in a clinical trial to be a 'what to represent' decision. This is because choice of comparator is *constitutive* of the scientific representation that the agents conducting the clinical trial intend to produce. To be clear: if agents conducting a clinical trial decide to compare adalimumab to methotrexate, *not* to sulfasalazine, then that decision *determines* that the trial will represent the relative efficacy of adalimumab and methotrexate— it determines that the trial will *not represent* the efficacy of sulfasalazine. For this reason, we argue, choice of comparator is a representational decision. So, too, we argue, are all decisions that are constitutive of a scientific representation that agents intend to produce: decisions concerning which study designs, categories, data sets, probability distributions, and parameter values to use are all representational decisions. As we will show, representational decisions have two important things in common: 1) they determine what information a scientific representation will include and exclude; 2) scientists' goal in making them is not to land on what is 'true', but rather on what is adequate for purpose.

To be sure, ours is not the usual way of understanding these decisions. The literature has yet to define a 'representational decision', and readers might question whether the decisions we characterize here as

representational are not better understood in more general terms related to the scientific process. Biddle and Kukla (2017), for example, in their foundational work on epistemic risk, use the term "phronetic risk" to describe "epistemic risks that arise *during the course of activities that are preconditions for or parts of empirical (inductive or abductive) reasoning*, insofar as these are risks that need to be managed and balanced in light of values and interests" (italics ours) (220). Readers may be inclined, thus far into our argument, to understand the decisions we highlight as 'phronetic' decisions. However, our aim is to encourage readers to recognize specific phronetic decisions *more precisely* as representational decisions, and thus to link the models in science and values in science literatures. With this link, it is clear that representational decisions are, in every important respect, central to science—yet impossible to make without values.

By now, philosophers widely agree that a type of scientific decision that is impossible to make without values is the type involving 'inductive risk' (Douglas 2000; 2009; Elliott and Richards 2017a). Here, we argue that the most fruitful definition of inductive risk is the risk of endorsing a 'fact' whose objective truth value is false. As we will show, this is Hempel's original conception, evident on a close reading of his (1935a; 1935b; 1954; 1965; 1981; 2000). We argue in favor of using Hempel's conception, as it helps us acknowledge the clear difference between endorsing a 'fact' and making a decision about representation. We then distinguish between 'hazards', 'hazardous events', and 'harms' within 'risk'[4] (Rausand 2011) and characterize representational decisions as a distinct hazard in science, which links to distinct hazardous events. We argue that this constitutes a unique and significant epistemic risk, which we call 'representational'.

In the models in science literature, a clear boundary already exists between a truth-apt claim and a representational tool. As we will show, drawing the same boundary in the values in science literature, and extending it to corresponding *decisions*, provides for an elegant integration of these literatures and their most deep-rooted insights. Perhaps most importantly, this integrated literature will establish that values cannot be limited to an 'indirect role' when informing representational decisions.[5] Representational decisions are normative decisions, requiring the same non-epistemic values that determine the purpose of an inquiry (cf. Peschard and van Fraassen 2014). Furthermore, representational decisions are the building blocks of scientific representations. These value-laden blocks do not build a value-free bridge between 'external' stages in science.

At the same time, models and other representational devices, at least in an ideal world, are only tools: whether to accept a claim that ostensibly follows from them should always be a further decision, one at a certain 'end' of the epistemic line. These 'end-line' decisions were a point of focus in early debates about values in science, as in, we show, Jeffrey's (1956) response to Rudner (1953). By distinguishing between inductive risk and representational risk, we can better appreciate the special character of these end-line decisions- and locate the precise (and unlikely) circumstances under which reporting probabilities, rather than accepting or rejecting hypotheses, can preserve the value-free ideal (VFI). In short, availing ourselves of the concept of representational risk will help us adjudicate central, historic debates in the values in science literature. Yet the concept of representational risk is not important only from an historical perspective, or to debates now (mostly) in the rear-view mirror. On the contrary, the concept of representational risk informs current problems at the crossing of values in science and social epistemology, including epistemic injustice in science and value entrenchment in complex, multi-authored models.

## 2. ADEQUACY FOR PURPOSE OR INDUCTIVE RISK?

We start with two home truths: models are not truth-apt but are evaluated for their adequacy for purpose- and *purposes reflect values*. What, then, is the epistemological significance of modelling decisions—or, more generally, decisions about scientific representation? These decisions have been sometimes framed as value judgments about the aims of research (Intemann 2015; Parker and Winsberg 2018), other times as having a subtle influence and significance distinct from inductive risk (Biddle and Winsberg 2009; Parker and Winsberg 2018; Winsberg 2010, 2012), other times as subject to inductive risk (Elliott and Richards 2017a; Parker and Lusk 2019; Steel 2015). Consider, first, different discussions of values in climate modelling, starting with threads where the term 'inductive risk' is avoided (Intemann 2015; Parker and Winsberg 2018; Winsberg 2010, 2012). In her 2015 work, Intemann submits that the goal of climate modelling is to generate "useful" predictions and, thus, value judgments must be made about the aims of research and "the extent to which particular practices, methodologies, or models are likely to promote those aims" (Intemann 2015, 219). In their own discussion of climate modelling, Parker and Winsberg (2018) acknowledge that modelling goals often stem from "non-epistemic interests and values" and that purposes and priorities determine what things are represented in a model and how they

are represented (128). Parker and Winsberg (2018) even establish that parameter values themselves may vary as a function of the priorities guiding calibration efforts, e.g., ensuring model output aligns with observations of particular variables (128). In this respect, they remark, modelling results are dependent on the non-epistemic values that determine purposes and priorities, as different values around these would produce different results (Parker and Winsberg 2018, 129). However, they concede that this point "is a variation on the familiar point that our current knowledge at any given time depends on (among other things) what we considered important enough to investigate, and thus on our interests and values" (Parker and Winsberg 2018, 130).

In our interpretation, the above remarks suggest that certain modelling decisions are viewed as value judgments associated with the 'external' phase of science (Bueter 2015; Douglas 2016; Elliott and McKaughan 2009). Parker and Winsberg (2018), at least, directly suggest that some modelling decisions equate to 'choice of question', an external value judgment of the important-yet-not-controversial variety (Longino 1990, 83-85; Elliott and McKaughan 2009). At the same time, both Intemann (2015) and Parker and Winsberg (2018) acknowledge that these decisions will influence model results. This reinforces the conclusion that 'external' value judgments have a more significant influence on theory appraisal than is generally appreciated (Bueter 2015; Elliott and McKaughan 2009; Winsberg 2018, Ch.9)—although neither Intemann (2015) nor Parker and Winsberg (2018) mounts a direct challenge to the internal/external distinction. Indeed, Parker and Winsberg (2018) seem to accept the premise that the value-dependency they initially identify is epistemologically insignificant. At least, they concede it exists, then go on to emphasize the (lesser-known) fact that estimates from a given model routinely become entrenched and incorporated into other models. Parker and Winsberg's (2018) argument echoes points made by Winsberg (2010, 2012), who stresses that model performance depends on the context in which it was developed, including decisions to prioritize certain predictive tasks over others. For Winsberg, "predictive preferences" (2012, 131) are distinct from inductive risk, yet epistemologically *significant* because they influence the estimation of uncertainties in a subtle way, beyond what can be tracked in practice or adjusted for in assigning probabilities to hypotheses (see Winsberg 2010, especially 110-111; 2012). Thus, for Parker and Winsberg (2018) and Winsberg (2010, 2012), value judgments about the aims of research are epistemologically insignificant *until* they exert an influence

beyond what is well-recognized—until they become embedded in the "nooks and crannies" of models (Winsberg 2012, 132).

Consider, now, a discussion of values in climate modelling centered around inductive risk (Parker and Lusk 2019). Parker and Lusk (2019) indicate that inductive risk is relevant to modelling decisions, particularly when "uncertain methodological choices are unavoidable" (1645) and there is a "risk of error" (1644). Following Douglas (2000, 2009), they explain that, on the inductive risk view, if it is ever unclear which methodological option "will give the most accurate results, scientists should consider how each option would affect the risk of different types of error and how bad the consequences of those errors would be" (1644). They write: "if choices [in climate modelling] must be made, they could be made in light of the inductive risk preferences of the user or client: if it would be particularly bad *for the user's purposes* for uncertainty to be underestimated, then the provider might select those methodological options that will deliver a broader uncertainty estimate" (Parker and Lusk 2019, 1647, italics ours). This raises an important question: if purposes and priorities determine what things are represented in a model and how they are represented, including *parameter values themselves and model results* (Parker and Winsberg 2018, 128), do purposes and priorities not present a risk of error? Parker and Lusk (2019) acknowledge that modelling decisions are made with the model user's purposes in mind. What is the difference between 'error' and inadequacy for purpose?

One ought to be able to turn to the definition of inductive risk to know how to distinguish it from decisions around adequacy for purpose. Although there are a few definitions of inductive risk (Elliott and Richards 2017b), by far the most influential is due to Douglas (2000, 2009). Douglas (2000) interprets Hempel's inductive risk as "the risk of error in accepting or rejecting hypotheses" (561), but she marks inductive risk as present in choosing a research methodology and gathering, characterizing, and interpreting data (565), which makes it clear that her own definition does not refer to hypothesis acceptance/rejection alone. Rather, Douglas' (2000) definition of inductive risk refers more broadly to the risk of error generally (559, 572). Critically, error is an ambiguous term, which can refer to any sort of mistake in judgment. Thus, it is necessary to read Douglas (2000) in detail for a full understanding of her definition. Most specifically, Douglas (2000) remarks:

> "...significant inductive risk is present at each of the three "internal" stages of science: choice of methodology, gathering and characterization of the

> data, and interpretation of the data. At each point, one can make a wrong (i.e., epistemically incorrect) choice, with consequences following from that choice. A chosen methodology assumed to be reliable may not be. A piece of data accepted as sound may be the product of error. An interpretation may rely on a selected background assumption that is erroneous. Thus, just as there is inductive risk for accepting theories, there is inductive risk for accepting methodologies, data, and interpretations." (565)

In our interpretation, Douglas's definition of 'error' is "a wrong (i.e., epistemically incorrect) choice" and her definition of inductive risk is the possibility of this type of error in the internal stages of science. Still, this definition requires some further analysis due to the use of the term "incorrect", which can mean *not correct as to fact* or, more generally, *inappropriate*. Arguably, Douglas (2000) does not mean 'not correct as to fact', since her definition locates inductive risk in choice of methodology, gathering and characterization of data, which include choices that, by broad consensus, cannot be 'factually' right or wrong. To avoid making a category mistake, Douglas (2000) must mean 'incorrect' in the sense of 'inappropriate'. However, this leaves us with the same confusion, since it is not clear what the difference is between inappropriate and inadequate for purpose.

Ultimately, Douglas (2000) does not give us the resources to distinguish 'error' from inadequacy for purpose. Rather, it seems, Douglas' (2000) inductive risk *is* the risk of making a choice that is inadequate for purpose, specifically *in the internal stages of science*. We argue that this definition is not as useful as it could be for understanding the role of values in science. For one, there is no 'factual' or 'essential' line between external and internal stages in science: it is up to philosophers to draw the line in whatever way serves their purpose. Thus, in practice, Douglas' (2000) definition of inductive risk is no more specific than 'the risk of making a choice that is inadequate for purpose'; the definition simply invites philosophers to defend (or not) their own interpretation of 'internal stages of science'. Already, philosophers have begun to analyze the 'inductive risk' in choosing to use composite outcome scores and specific clinical trial designs (Bluhm 2017; Stanev 2017). In our view, either the next step is for philosophers to debate whether such choices 'really occur' in the 'internal' stages of science, or to decide whether they really have a special interest in those stages as far they understand them now. Elliott and McKaughan (2009), Bueter (2015) and Winsberg (2018, Ch. 9) have already given us good reasons to have an *equal interest* in the *external* stages of science, including

the fact that decisions made in these stages influence *which hypotheses are accepted and rejected*. As Winsberg (2018) puts it, "deciding what data to collect, and when to stop collecting more, can have a strong influence on what probabilities we will assign to hypotheses" (137). To illustrate his point, Winsberg asks us to imagine that it is 1925 and the hypothesis that smoking causes cancer is generally being assigned a low probability. Should society spend money to study whether the hypothesis is true? Winsberg remarks:

> "The answer to the question of whether we should spend the money will have an obvious impact on the degree of belief we assign to the hypothesis. Of course the 1925 scientists cannot predict whether the research will raise or lower the probability of the hypothesis, but they can predict that it will very likely push that probability in one direction or another. This blurs the line between the epistemic and the normative in a way that most commenters seem to believe won't happen if the values only play a role that is "external" to science." (Winsberg 2018, 137)

This conclusion, we think, is a sufficient reason to set aside the external/internal distinction and seek a more useful definition of inductive risk that does not incorporate it. In the remainder of this paper, we will argue that the more useful distinction to draw is between endorsing a 'fact' and making a representational decision. In turn, we will argue in favor of defining inductive risk as the *risk of endorsing a 'fact' whose objective truth value is false*. This, as we will show, is how Hempel defined it, and putting this historical definition to use has numerous advantages. Among them is distinguishing inductive risk from 'representational risk', the risk of making a *representational decision that is inadequate for purpose*. As we will show, this risk is not only epistemically interesting, but philosophically fruitful in a number of ways.

## 3. HEMPEL'S INDUCTIVE RISK

To interpret Hempel's account of inductive risk, it is useful to trace its roots in debates in early 20th Century, logical positivist discussions of confirmation and empirical significance. Readers will recall that many logical positivists believed, at least early on, that in order to be meaningful, a synthetic claim must be implied by a finite number of observation sentences. Many different versions of this claim, including many weakenings of it, were defended, but always central was the idea that meaningful discourse was tightly connected to confirmation, where

confirmation was a relation that obtained between various kinds of claims and a special kind of claim: observation statements, or what members of the Vienna Circle often called 'protocol statements'.

Hempel (1935a) reviews Carnap's early conception of protocol statements ("the result of pure immediate experience without any theoretical addition") and distinction between empirical laws ("general implicative statements") and singular statements (e.g., "Here is now a temperature of 20 degrees centigrade") (51-52). On this view, general statements are tested by their singular consequences, i.e., by unique experimental or experiential results. However, because singular consequences are infinite, general statements can never be fully verified, only more or less supported. This prompts Hempel to conclude that all general statements have the character of a hypothesis (1935a, 52). He then recalls Carnap's own conclusion that singular statements have the same character: singular statements (like "here is now a black raven") can only be more or less confirmed, and there is no clear minimum degree of confirmation for a singular statement to be adopted; thus, the adoption or rejection of a singular statement depends upon a decision (1935a, 58). These decisions are at the heart of Hempel's inductive risk.

Turning to interpret Carnap's and Neurath's more recent ideas, Hempel (1935a) comments: "Even the protocol statements are revealed to be hypotheses in relation to other statements of the whole system; and so a protocol statement, like every other statement, is at the end *adopted or rejected by a decision*" (58, italics ours). This implies that all empirical statements have the character of a hypothesis. The same view is evident in Hempel's 1935 essay (1935b , 96) and 1937 essay "The Problem of Truth" (Hempel 2000 [1937], 52).[6] Thus, for Hempel, deciding to adopt any empirical statement counts as accepting or rejecting a hypothesis. Adopting "all ravens are black" and "this is a black raven" or even, "this thing in front of me is black" are all cases of adopting a hypothesis and all involve a decision.

In introducing the term 'inductive risk' in his 1954 essay, Hempel uses the example of attributing the disposition 'solubility-in-water' to a lump of sugar that is not actually put in water (Hempel 1954). He remarks that to do so "is to make a generalization, and this involves an inductive risk" (14), further noting that if we were to reject any procedure that involves inductive risk we would be prohibited from using dispositional concepts (14-15). In revisiting the concept of inductive risk in his (1965), Hempel notes that any empirical law is accepted on the basis of incomplete

evidence, though "[such] an acceptance carries with it the 'inductive risk' that the presumptive law does not hold in full generality" (92). While his ensuing discussion of statistical testing might suggest that his remarks on values apply only to hypothesis acceptance or rejection in that context, we do not interpret them so narrowly. Rather, we take it that Hempel uses examples from statistical testing because they are salient and rhetorically effective, though he is aware his comments are more far-reaching. Hempel confirms this in his essay "Turns in the Evolution of the Problem of Induction" (1981 ), in which he discusses the debate between Rudner (1953) and Jeffrey (1956). To a list of difficulties facing Jeffrey's argument Hempel adds:

> "Even if the scientist limits himself to determining probabilities for hypotheses, he must perform tests to obtain the evidence on the basis of which to calculate those probabilities. He must, therefore, it seems, accept certain empirical statements after all, namely *the evidence sentences by which he judges the probability of contemplated hypotheses*."[7] (Hempel 1981, 395, italics ours)

Thus, though Hempel in the above uses 'hypotheses' in the limited sense in which frequentists nowadays use the term, he confirms that his inductive risk extends to the acceptance of empirical statements generally. This is consistent, we note, with Rudner's (1953) conclusion. Although Rudner is sometimes read as suggesting that scientists make value judgments only at the end of a formal hypothesis test or empirical study, at the end of his essay he explicitly extends his conclusions to all scientific hypotheses, in the same sense as Hempel. Rudner (1953, 5) quotes Quine (1951):

> "Science is a unified structure, and in principle it is the structure as a whole, and not its component statements one by one, that experience confirms or shows to be imperfect. Carnap maintains that ontological questions, and likewise questions of logical or mathematical principle, are questions not of fact but of choosing a convenient conceptual scheme or framework for science; and with this I agree only if the same be conceded for every scientific hypothesis." (Quine 1951, 72)

Readers will recall that Quine, pace Carnap, argued that there was no clear line between internal and external questions, and hence that if Carnap was right that a choice of logic, for example, involved a value judgment, then so did the choice to say that 'there is a black swan here' (Leitgeb and Carus 2020). Thus, while we think the Quinean Rudner would disagree with Hempel on the reasons why it is so, he would agree

that all statements of matters of fact, whether statistical or not, involve a value judgment.

This brings us to consider how best to define Hempel's inductive risk. For clarity, we conform our definition to the bow-tie model of risk, i.e., risk as hazard, hazardous event, and harm (Rausand 2011):

*Hempel's Inductive Risk (IR)*

**Hazard:** endorsing a 'fact', a truth-apt claim about the world with an objective truth value[8]
**Hazardous event:** endorsing a 'fact' whose objective truth value is false
**Harm:** whatever undesired consequences, specified or unspecified, may follow from endorsing the 'false fact'

In what follows, we will contrast Hempel's inductive risk with 'representational risk', which we will define as follows:

*Representational Risk (RR)*

**Hazard:** making a representational decision in science
**Hazardous event:** making a representational decision that is inadequate for purpose, either for our own purpose or for the purpose of other epistemic agents
**Harm:** whatever undesired consequences, specified or unspecified, may follow from making an inadequate representational decision

As we will show in the next section, distinguishing between IR and RR helps clarify the significance of Jeffrey's (1956) response to Rudner, and allows us to acknowledge certain differences between rationalist and personalist probabilists. After establishing this, we go on to show just how epistemically interesting it is to compare IR and RR in terms of the hazardous events and harms that may result from each, and to consider the background conditions that are relevant to RR.

## 4. WHAT'S THE RISK IN REPORTING PROBABILITIES?

Readers will recall that Jeffrey (1956) replies to Rudner (1953) that it is not the job of the scientist *qua* scientist to accept or reject hypotheses; rather, she must only assign probabilities to hypotheses. Here, we apply the distinct concepts of IR and RR in order to establish what risk the scientist *qua* scientist faces in assigning probabilities to hypotheses and what implications this has for the VFI. In doing so, we defend our interpretation of Jeffrey's (1956) response to Rudner (1953).

Philosophers sometimes take Jeffrey to be claiming that he has rebutted Rudner regarding the untenability of the VFI—more specifically, to be claiming that the VFI *is* tenable because scientists need not accept or reject hypotheses.[9] However, a close reading of Jeffrey's two-paragraph conclusion shows that he is rebutting in Rudner just the claim about scientists' being obliged to accept or reject hypotheses—and that, in fact, he marks the VFI's tenability as an open question that depends on the result of debates in the foundations of probability (Jeffrey 1956, 245-246). Ending the first paragraph of his conclusion, Jeffrey writes, "In any event, we conclude that it is not the business of the scientist as such, least of all of the scientist who works with lawlike hypotheses[10], to accept or reject hypotheses" (245). Opening the second he writes, "We *seem* to have been driven to the conclusion that the scientist's proper role is to provide the rational agents in the society which he represents with probabilities for the hypotheses which on the other account he simply accepts or rejects" (245, italics ours). Here, we might take Jeffrey as thinking he has saved the VFI, if we were to ignore that he goes on to list *"great difficulties with this view"*, including that it "presupposes a satisfactory theory of probability in the sense of degree of confirmation for hypotheses on given evidence" (245, italics ours).

What Jeffrey refers to here is a logic of confirmation, something to make it a matter of logic what the probability of a hypothesis is conditional on the evidence, as was sought by the Carnapian rationalist probabilists. Readers will recall that there are two competing schools of thought among 'probabilists' about the nature of 'original' prior probabilities (Galavotti 1991). According to the first, 'rationalist' school (represented by Carnap), there is an *objectively correct* probability that a scientist should attach to each hypothesis in light of the evidence she has. According to the second, 'personalist' school (represented by Ramsey and de Finetti), any set of priors whatsoever is allowable, as long as it obeys the axioms of probability. On the personalist view, probabilities are simply things that we have; they are neither things we decide on, *nor things that are objectively true or false*. Having defined the Hempel's IR above, we can say that, for the rationalist probabilist, there is IR[11] in attaching probabilities to hypotheses, but for the personalist, there is not. This is because the rationalist and the personalist construe probabilities in fundamentally different ways, and only on the rationalist view does reporting a probability amount to endorsing an objectively true or false 'fact'.

In referring to the search for a logic of confirmation, Jeffrey (himself a thinker more often associated with personalist probabilism) makes an important point. That is: to save the VFI, not only must the job of scientists be limited to providing probabilities, but the rationalists must be fully successful in their project. That is, for scientists to *provide* probabilities *in a value free manner* a) the rationalists must be right that there is a unique correct logic of probability and b) scientists must have it in hand and be omniscient with respect to it.

Let us now compare the consequences of the (various) rationalist and personalist probabilist positions for what hazards scientists engage in, what hazardous events scientists risk, and what harms scientists can bring about. Again, for the rationalist, there is an objectively correct probability the scientist should attach to each hypothesis in light of the evidence she has. This can free her from engaging in any hazard *if* there is a uniquely correct logic of probabilities, and she is a perfect reasoner with it. Reporting the probability she calculates, then, will present *no hazard* since there will be *no possibility of being wrong*. However, a hazard will present itself if:

1. There is a uniquely correct logic but the scientist is not a perfect reasoner with it
2. There is more than one allowable logic (as Carnap maintained)

In the first case, the scientist will be making a decision that can be objectively wrong, and so a possible hazardous event is making an objectively wrong claim. The corresponding risk is Hempel's IR. In the second case, the hazard is the choice of which allowable logic to use. Here, there is (by stipulation) no objectively correct logic, so the possible hazardous event is not choosing an objectively wrong one. It follows that there is no IR. However, a possible hazardous event is choosing a suboptimal logic. Choosing a logic is a representational decision, and, at least according to Carnap, a logic can only be suboptimal with respect to some pragmatic framework. Given the possibility that the scientist will choose a logic that is suboptimal with respect to her own, or her stakeholders', pragmatic goals, she faces RR.

Turning now to the personalists, readers will recall three key features of their position. First, any set of priors is allowed, as long as it obeys the axioms of probability. Second, probabilities are not things we decide on, nor things that are objectively true or false, but simply degrees of belief. Third, degrees of belief are cognitive states that, in concert with non-cognitive states like utilities, guide all behavior. While 'having' probabilities

is not a behavior, *providing probabilities* is one indeed, and the scientist will provide to her 'consumer' whatever probability the providing of which she thinks will maximize her own expected utility.[12]

It follows from the above that, for the personalist, there is no IR in reporting probabilities, not in 'having' them, nor in 'providing' them. However, providing probabilities involves representation, and there is a possible hazardous event in this process: possibly, the scientist will *provide probabilities that are based on the wrong utilities*—either utilities the scientist later comes to regret, or utilities that do not represent those of her stakeholders. So while personalism about probability frees scientists from IR, it does not free them from RR, or secure the VFI, as long as they are engaged in providing probabilities.

## 5. REPRESENTATIONAL RISK

Within the risk concept, hazards, hazardous events, and harms may be distinguished (Rausand 2011).[13] A banana peel in one's book bag does not lead to slipping; a banana peel on the floor does not lead to a smelly book bag. A slipper may dust himself off or head to the hospital; a book bag may or may not need to be replaced altogether—a number of background conditions may determine the ultimate outcome.[14] In the previous section, we contrasted two hazards—accepting/rejecting a hypothesis and providing probabilities—and showed that personalists will regard these as distinct and appreciate that each leads to a unique hazardous event (even if downstream harms are left unspecified). In this section, we review representational decisions *other* than providing probabilities, and further illustrate how they differ from endorsing a 'fact'. Once there can be no doubt that representational decisions are a distinct hazard, risking a distinct hazardous event, we turn to specifying some of their potential harms under different background conditions relevant to RR.

### *5.1 Distinct Hazards, Distinct Hazardous Events*

In Hempel's IR, the hazard is endorsing a 'fact', a truth-apt claim about the world with an objective truth value. With respect to some representational decisions, it is easy to see that Hempel's hazard is not at issue, as nothing resembling an objective truth value is anywhere to be found. For example, scientists do not speak of a 'false' study design; none argues that composite outcome measures are 'true'. Diagnostic criteria for disease may shift with a growing understanding of facts, but those facts do not determine what the disease is *truly*. It is not in these cases that there

may still be some doubt that making representational decisions differs from endorsing facts. Rather, we think, this doubt might linger when it comes to representational decisions that concern a quantitative value, such as parameters in simulation models. For example, it might seem that if a climate model inputs the acceleration of gravity as 9.8 $m/s^2$ it is because that is the true value. Indeed, in this context, modellers may seek—and find—something that resembles an objective truth value. However, this is a red herring: where model parameters are populated with 'true' values, those values are not chosen *because* they are true, but because they are adequate for purpose. This is made clear by the fact that, in many cases, only a *false* parameter value may be adequate (Parker 2010, 990). For example, a climate model might make the acceleration of gravity vary from the equator to the poles to adjust for its exclusion of the centrifugal 'force' associated with the Earth's rotation. Or consider parameter values for cloud formation in climate models. Here, modellers do not choose the value that best represents cloud formation, but the one that best balances out the fact that all climate models leak energy at the top of the atmosphere (Winsberg 2018, Ch.10). Furthermore, a great many quantitative values input to representational tools do not have a corresponding truth value. To use Parker's (2010, 990) example, there is no single speed at which all ice crystals fall in clouds. To use Winsberg's (2010, 82) example, there is no true value for the bonding energy of a "silogen atom", since there is no such thing (despite the fact that they are modelled in crack propagation).

From the above examples, we should draw a broad lesson: making a representational decision differs from endorsing a fact. They are distinct hazards: the latter leads to a very specific hazardous event (endorsing a 'fact' whose objective truth value is false), the former a more variegated inadequacy. Still, a sticking point remains: in some models, the 'true' value for a parameter will be the only one considered adequate for purpose. In this case, it seems, inadequacy would constitute falseness. This may invite the objection that populating these values involves *inductive* risk, blurring the purported distinction between it and RR. However, readers familiar with sensitivity analyses will continue to appreciate the distinction between these two risks. In models where having the 'true' value for a parameter is of great importance, it is common to input a range of values for that parameter to see the effects of getting it wrong. Sometimes, modellers will find that the 'true' value can be varied to some, even great, extent without affecting model results. In other words, there will be no downstream harm that results from incorporating a 'false fact' into the representation. In

this case, IR is not an apt descriptor. Other times, results are found to be sensitive to that value, and modellers face a subsequent decision regarding *how to represent* their findings, including justifying the range of values used in sensitivity analyses (after all, this range typically has no corresponding truth value). In this case, IR is not an apt descriptor either. The value of distinguishing between IR and RR will become even clearer when we consider the unique harms that can result from representational decisions, which extend beyond those that stem from endorsing a 'false fact'.

One last sticking point: a representational decision that is inadequate for purpose can most certainly lead to an inference to a 'false fact'. For example, Stanev (2017) describes how the use of a composite outcome measure in a clinical trial resulted in an inference to a 'false fact' about a drug's effects. In this context, readers might object, the 'hazardous event' in our so-called RR is identical to that in Hempel's IR. Our reply is that this is not quite right: rather, through the IR lens, the hazardous event is an inference to a 'false fact', and through the RR lens, the hazardous event is inadequacy for purpose. This matters, we argue, because the RR lens invites us to view an inference to a false fact as a *downstream harm*, which emphasizes that it can be *prevented upstream*. For example, in analyzing Stanev's (2017) case study through the RR lens, we are encouraged to ask under what circumstances a composite outcome score is adequate for purpose, and what is necessary to ensure one does not persuade us to endorse a 'false fact'. With this, we turn to how representational decisions can result in harms unrelated to endorsing a 'false fact', under a variety of background conditions.

### *5.2 Distinct Harms Under Distinct Conditions*

In general, representational decisions determine what information will be highlighted and what information will be obscured by a scientific representation (van Fraassen 2008) . For a representational decision to be adequate for purpose, it must highlight the information that the (relevant) epistemic agents desire, and obscure only the information that they do not. When representational decisions are inadequate for purpose, one possible downstream event is that epistemic agents will endorse a 'false fact' and harm will result from this. However, representational decisions can result in harms that are unrelated to 'false facts': in lamentably incomplete scientific results, irrelevant or distracting results, even pernicious and unjust gaps in scientific knowledge. The concept of RR, we argue, invites us to consider these distinct harms, and the role of background conditions in determining

whether or not they will result from representational decisions. We show this by considering well-known issues in the design of clinical trials, and their influence on trial results and (later on) simulation models in health economics.

Consider, first, the choice of comparators (i.e., experimental and control interventions) in a clinical trial. This representational decision is a straightforward extension of the research question: for example, a clinical trial may aim to answer whether an experimental treatment is more efficacious than placebo, or more efficacious than an active treatment, such as the one typically used in clinical practice (the 'standard of care') (Anderson 2006). At the same time, choice of comparators has an obvious, significant influence on the content and impact of trial results (Bluhm 2017; Glasser and Howard 2006; Mann and Djulbegovic 2013; Wilholt 2009). One specific problem is the possibility of a "substandard comparison" (Wilholt 2009, 93) or "comparator bias" (Mann and Djulbegovic 2013, 30). These terms refer to the practice of comparing experimental treatments to control interventions that are less effective than the standard of care: either the "inappropriate" use of placebo or the use of an "inappropriate" active comparator (e.g., the standard of care in a suboptimal dose) (Mann and Djulbegovic (2013, 30–31). As Bluhm (2017) explains, many experts argue that using an appropriate comparator is an "epistemological requirement" (203):

> "This is because a trial should provide *knowledge that is useful* to those clinicians who would be using the results of a trial. What clinicians—and for that matter patients —want to know about a promising new medication is not whether it is better than a placebo, but whether it is a better therapy (or at least as good a therapy) as the one(s) already available and used in clinical practice. The only way to answer this question is to actually test the new drug against a current therapy." (Bluhm 2017, 203, italics ours)

The expectation that clinical trials should provide "knowledge that is useful" is an important condition: it indicates what is required for the choice of comparator to be adequate for purpose. Yet, obviously, different stakeholders have different views of what knowledge is useful. As Mann and Djulbegovic (2013) remark "It is not surprising, therefore, that researchers, sponsors, patients and government regulators may have different views on the selection of comparators" (106). It follows that to be able to define 'comparator bias'—to identify the use of an 'inappropriate' comparator—it is necessary to know *whose purposes* are meant to be fulfilled by a clinical trial. This points us to one important background

condition relevant to RR: for *whose purposes* a representational decision is expected to be adequate.

For the sake of argument, let us assume, as Bluhm suggests, that a clinical trial should provide knowledge that is useful to patients and clinicians. That is, that representational decisions, like choice of comparators, should be adequate for *their* purposes. If representational decisions in clinical trial design do not meet this expectation, *what is the harm?* (To be sure, it is not necessarily a 'false fact': clearly it is possible to provide information that it is 'true' but not the useful knowledge sought.) At a minimum, we think, the harm would be trial results that are lamentably incomplete from the perspective of patients and clinicians. At least, when a trial shows that a "promising new medication" is better than placebo—but not whether it is a better than a medication already used in clinical practice—we might call those results 'incomplete' and aim to further characterize the harm in that. At the same time, it seems that other, possibly more severe, harms might result from an inadequate choice of comparators under other conditions. For example, what if the experimental treatment under study is "promising" only from the perspective of the developer, and *not promising* from the perspective of patients and clinicians? What if, for example, the experimental treatment is a 'me-too' drug almost indistinguishable from an existing generic (Gastala et al. 2016), or a preventive medication for which there is little evidence of demand, even some evidence of disdain (Mosor et al. 2020)? In these cases, we argue, the representational decision has a built-in inadequacy, but the harm is not 'incomplete' results; rather, the harm is, at least, 'irrelevant' or 'distracting' results. Yet to distinguish between these subtly different harms requires additional knowledge of background conditions, of specific details about the *context of purposes* among clinicians and patients.

As much as scientific results that are incomplete, irrelevant, or distracting are clearly distinct from 'false' ones, one possible objection is that these are not really serious harms. Although we think this objection might stand under some conditions, it is unlikely to stand under others. Consider the significance, now, of choice of comparators not in a single clinical trial, but in clinical trials *generally*, specifically in the context of their overwhelming sponsorship by pharmaceutical companies (Lundh et al. 2017). Light and Lexchin (2020), for example, argue that clinical trials sponsored by pharmaceutical companies employ a "strategic ignorance": not only strategic choices of comparators, but numerous other strategies designed to hide the risk of adverse effects associated with experimental

treatments (4). These strategies include limiting the length of clinical trials, recording only certain adverse effects, and excluding participants with co-morbidities (Light and Lexchin 2020, 4), all of which, we note, relate to representational decisions and their capacity to obscure information. In our interpretation, what Light and Lexchin (2020) describe is a *systematic* and *persistent* obscuring of information by an institution of social power, which suggests that the resulting gaps in knowledge could potentially be viewed as an epistemic injustice (Fricker 2007; Kidd, Pohlhaus, and Medina 2017). This invites us to consider the *social power dynamics* that determine whose purposes inform representational decisions in clinical trials.

As Glasser and Howard (2006) emphasize, there are at least ten different issues in clinical trial design that can affect the outcome of a trial: choice of comparator, eligibility criteria, and selection of end points are just some examples of what we call representational decisions. With respect to these decisions, "There is no correct answer!" (Glasser and Howard 2006, 1108) other than the one that is adequate for purpose—and purposes differ among epistemic agents. This adds a layer of complexity to RR, since epistemic agents routinely pick up and use representational devices built by others, potentially with different purposes in mind. Just one form of this sort of 'secondary use' of clinical trial results occurs in health economics, where simulation models are used to compare the cost and effectiveness of alternative clinical interventions in order to inform health policy (Drummond et al., 2005). To build these models, health economists routinely derive 'effectiveness' parameter values from published clinical trials. In theory, more than one sort of inadequacy for purpose could arise in this context: for example, if participants with co-morbidities were excluded from a particular trial, the effectiveness estimate derived from that trial could be inadequate for purpose, as it may not generalize to the larger population of patients affected by health policy.[15] How do health economists determine if parameter values from published clinical trials are adequate for *their* purposes?

The secondary use of clinical trial results in health economics models, we think, points to a number of background conditions relevant to RR. One of the more obvious ones is individual and group-level *knowledge*: if health economists building a simulation model are *unaware* of the inadequacy of a parameter value for their purpose, surely they will be less likely to seek a more adequate one, or conduct sensitivity analyses to explore the effect of varying it. However, another important condition might well be epistemic complexity: as Parker and Winsberg (2018)

emphasize, modellers sometimes lack the *ability* to change a parameter value (or other representational choice), for multiple reasons rooted in the complexity of computational models. Still another might be the modellers' institutional context and the demands it makes (cf. Winsberg, Huebner, and Kukla 2014). For example, health economists sometimes make representational choices under time constraints, which may give an incentive to carry over representational choices from published models, even if they may be inadequate for purpose. This was established in a qualitative study of health economists, who were briefed about Parker and Winsberg's (2018) argument and asked about the possibility for values in modelling to have a 'cascade' effect (Harvard, Werker, and Silva 2020). One participant's description of building a model, we think, evokes both *epistemic complexity* and *institutional context*:

> "… me and another guy made it up in a week. That and the assumptions, literally in a week. A big week, because we were on a deadline … I still see people today using those same assumptions that we made in a week, today, in l like [distant country]. Are they the right assumptions? I don't know. They seem alright, but like, you see how that cascade happens." (Harvard, Werker, and Silva 2020, 9).

## 6. CONCLUSIONS AND IMPLICATIONS FOR CURRENT DEBATES IN VALUES IN SCIENCE

In this paper, we have defended three closely related proposals: to employ Hempel's conception of inductive risk; to draw a conceptual boundary between endorsing a 'fact' and making a representational decision; and to conceptualize 'representational risk' as a unique epistemic risk. In this section, we outline the implications of each proposal for current debates in the values in science literature.

In their epilogue to a collection of case studies in values in science, Elliott and Richards (2017b) highlight the ongoing debate over how to conceptualize inductive risk. Major areas of investigation for philosophers, they remark, include questions about the "nature" of inductive risk, like "Must inductive risk involve the acceptance or rejection of hypotheses?" and "Should the argument from inductive risk focus on errors or on standards of evidence?" (Elliott and Richards 2017b, 274). Here, we have argued in favor of employing Hempel's conception of inductive risk: the risk of endorsing a 'fact' whose objective truth value is false. This conception, we note, gives us straightforward answers to the above

questions: that is, inductive risk must involve the acceptance or rejection of a truth-apt claim about the world with an objective truth value, and the AIR should focus on 'errors' in the specific sense of endorsing a 'fact' whose objective truth value is false. Answering these two questions is advantageous, particularly because conceptual clarity around inductive risk, specifically, should help advance the literature on "epistemic risk" generally (Biddle and Kukla 2017; Biddle 2018). Biddle and Kukla (2017), for example, argue persuasively that not all risks in science should be construed as inductive risk; and many philosophers appear sympathetic to this idea, as the term epistemic risk has begun to appear in the values in science literature (e.g., Valles 2018; Winsberg, Oreskes, and Lloyd 2020). However, Biddle and Kukla's (2017) argument is weakened somewhat by a degree of ambiguity in their definition of inductive risk. This ambiguity is noticeable in a passage in which they associate inductive risk both with inference from evidence, generally, and inference from statistical evidence, specifically:

> "One variety of epistemic risk, as we have already seen, is inductive risk — again, traditionally understood as the risk of wrongly accepting or rejecting a hypothesis on the basis of evidence. Inductive risk is in at least one important sense different in kind from alethic risk, as it is located at a certain point during the practical process of settling our beliefs and generating knowledge, *namely in the inference from statistical evidence to an empirical conclusion.*" (Biddle and Kukla 2017, 218, italics ours).

Biddle, too, sometimes gives the impression that he associates inductive risk specifically with inferences from statistical evidence (2018, 2020a), other times with evidence generally (2020b). For example, in analyzing two cases of epistemic risk, he comments "Neither of these risks is best thought of as inductive risk; *neither is a mistake made in inferring a hypothesis from statistical evidence*" (Biddle 2018, 365, italics ours). We argue that employing Hempel's conception of inductive risk will eliminate any ambiguity around this core concept, and better position philosophers to distinguish inductive risk from other varieties of epistemic risk. This, we argue, is significant, as we agree with Biddle (2018) that it is important "to investigate more fully the various types of epistemic risk, how they relate to one another, and how they might be managed" (365). We will return to the concept of epistemic risk once more below.

Other debates in the inductive risk literature concern the distinction between 'direct' and 'indirect' roles for values in science. Specifically, Elliott

and Richards (2017b, 274) ask "How is the distinction between direct and indirect roles best characterized in practice?", and "Is the distinction between direct and indirect roles the best replacement for the VFI?", marking these as major questions to address in future investigations of inductive risk. In short, our answer to the first question is: with further reference to the difference between endorsing a 'fact' and making a representational decision. Indeed, a lack of attention to this difference seems to contribute to the ambiguity in Douglas' (2009) characterization of the direct/indirect role distinction (cf. Elliott 2013). As Elliott (2013) shows, Douglas develops two interpretations of the distinction: on the 'Logical' distinction, values can be treated as warrant or evidence *for a claim* (direct role), or values can influence decisions about how much evidence is sufficient to accept a claim (indirect role) (Elliott 2013, 377). On the 'Consequential' distinction, values can influence scientists' choices based on intended outcomes that they want to bring about by accepting a claim (direct role) or values can influence scientists' choices based on unintended consequences associated with mistakes that they want to avoid (indirect role) (Elliott 2013, 377). To develop an unambiguous distinction between direct/indirect roles for values, we think, it is essential to define both the terms 'claim' and 'outcome'[16] with reference to representation. For one, as we have shown, representational decisions do not amount to 'claims'. Furthermore, outcomes of representational decisions (i.e., outcomes manifest in representations) differ from outcomes of decisions whether or not to endorse a fact (i.e., outcomes manifest in the actions of epistemic agents). One of the key benefits of distinguishing between endorsing a 'fact' and making a representational decision is seeing that *representation specifically* puts non-epistemic values in a 'direct' role—that is, a straightforward and conspicuous one. As we have shown, the purpose of a representation always informs representational decisions: nothing makes this clearer than legitimate, purposeful choices of false parameter values in simulation models.

Our answer to Elliott and Richards' first question, we think, should also help inform their second, i.e., whether the direct/indirect role distinction is the "best replacement for the VFI" (Elliott and Richards 2017b, 274). Many philosophers seem to think the direct/indirect role distinction cannot be the 'best replacement' because it tries, yet fails, to address the problem of wishful thinking (Elliott 2013; de Melo-Martín and Intemann 2016; Steel and Whyte 2012). As de Melo-Martín and Intemann (2016) put it, the direct/indirect role distinction tries to address the problem of wishful

thinking by ensuring that values "play no direct role in *determining what the evidence is*, such that a predetermined outcome would be favored" (509, italics ours), yet fails because decisions about how much evidence is needed "can indeed influence what evidence there is, or "rig" the methodologies used toward achieving a predesired outcome" (509). This seems right: for the direct/indirect role distinction to be a good replacement for the VFI, it would need to address the problem of wishful thinking. The distinction between RR and IR points us to one way the fundamental proposal might be salvageable. One possibility worth evaluating, in other words, is that non-epistemic values should be recognized as inevitably playing a direct role in informing representational decisions, but should be limited to an indirect role in informing decisions to endorse a 'fact'.

Our last proposal is to conceptualize representational risk as a unique epistemic risk. Indeed, we wish to argue that representational risk should be regarded as a 'core' type of epistemic risk, equal in significance to inductive risk. In our view, representational decisions and decisions to endorse a 'fact' are both core hazards in science—on par with one another at least in terms of incidence and prevalence—yet ones that should be distinguished and analyzed separately for the risks they create. Our argument, we think, advances the literature on epistemic risk by addressing two of its current shortcomings (Biddle and Kukla 2017; Biddle 2018, 2020a, 2020b). First, most modestly, the epistemic risk literature lacks a framework to assist in distinguishing between epistemic risks, and we think our method of contrasting hazards, hazardous events, harms, and background conditions will prove useful in this respect. Second, and more importantly, Biddle and Kukla (2017) and Biddle (2018) have argued that inductive risk is just one type of epistemic risk and that there are "many others" (e.g., Biddle 2018, 365), yet their identification of epistemic risks has been more pragmatic than systematic. At the very least, they have not sought to establish whether any epistemic risks rival inductive risk (e.g., in terms of their incidence and prevalence in science), or to analyze in detail the relationships between the different epistemic risks they identify. We note that many of the "phronetic" risks that Biddle and Kukla (2017) identify—data formation risk, model choice risk, conceptual definition and operationalization risk (222), power risk, framing risk (Biddle 2018)—involve representational decisions. By conceptualizing representational risk, we aim to establish that there is indeed an epistemic risk that rivals inductive risk, and that many of the epistemic risks identified so far appear to be under its umbrella.

NOTES

1. Following Longino (1990, 85-86), Douglas (2000, 2009, 2016) uses the term "external" to describe the phases (stages, parts) of science in which non-epistemic value judgments are well-recognized and non-controversial; Winsberg (2018, Ch.9) also uses the term "external", while Elliott and McKaughan (2009) refer to "discovery and pursuit" and Bueter (2015) refers to the "context of discovery". We take all of the above terms to include "the choice of areas or aspects of the world to be illuminated by application of the rules [of scientific inquiry]" (Longino 1990, 85) and refer to value judgments that philosophers of science generally regard as "epistemically uninteresting" (Elliott and McKaughan 2009, 600). For brevity, throughout this paper, we use the terms 'external' and 'internal' to the distinguish between phases of science widely taken to involve uncontroversial/uninteresting value judgments and controversial/interesting ones, respectively.
2. Throughout this paper, we modify Hempel's phrase 'adopting' (also 'accepting') a 'fact' to 'endorsing a fact', which we take to mean *endorsing a truth-apt claim about mind independent objects*. We use scare quotes around 'fact' to indicate that we mean a claim of facticity, not that the claim is necessarily true. We replace 'adopt' (and 'accept') with 'endorse' because we take inductive risk, linked as it is to decision theory, to be associated with a choice of action, rather than with a cognitive attitude (the latter of which might or might not be conceptualized as a free choice). This is also in line with our discussion of probabilities in Section 4, where we take the central issue to be one of 'reporting' a probability, rather than having one. We do not take a view on whether Hempel considered the distinction between taking an action and adopting a cognitive attitude.
3. An important exception is Peschard and van Fraassen (2014), who directly discuss the normativity of 'relevance judgments' in experimental modelling, noting "How we view modeling today gives reason to expect some novel insight in the role of values in science there" (4). In our interpretation, Alexandrova (2010) and Parker (2010, 2020) make it clear modelling is a value-laden process but do not draw explicit connections to the values in science literature.
4. Following Rausand (2011), risk involves a **hazard** (e.g., riding a motorcycle), a **hazardous event** (e.g., crashing), and that event's **undesired consequences or harm** (e.g., getting injured or killed).
5. Douglas (2009, especially Ch. 5) argues that non-epistemic values can legitimately play an 'indirect' role throughout scientific inquiry (i.e., in assessing the consequences of error and determining evidential standards) but should

play a 'direct' role only in the early stages of scientific inquiry. Elliott (2013) argues that Douglas' (2009) proposal is both ambiguous and complicated by the multiple goals of science; we build here on his latter point. A related debate (in our interpretation) concerns whether it is possible for epistemic values to take 'priority' in scientific inquiry (e.g., Brown 2017; Steel 2017). As we will make clear, our position is that 'epistemic priority' is not possible throughout the process of representation.

6. e.g., "...any of the ordinary physical statements, even such as 'This is a piece of iron,' is a hypothesis the adoption of which depends in the end upon a convention." (Hempel 1935b, 96); "...this system of observation sentences does not form an absolutely fixed and unshakeable basis, for each directly verifiable sentence is itself only a hypothesis." (Hempel 2000, 52).
7. Of course, Jeffrey might reply that even the evidence sentences can be assigned probabilities, rather than accepted or rejected. For now, it suffices to take this passage to be textual evidence that Hempel included all empirical statements, even basic ones, under the umbrella of hypotheses about which there is inductive risk.
8. An anonymous referee reminds us that much of the literature defines inductive risk so that it is associated not only with endorsing a fact that is false, but *failing to endorse* a fact that is true. Here, we should note that *failing to endorse a fact* comprises two alternatives: endorsing its negation and remaining silent. It is clear that the risk involved in endorsing its negation is Hempel's IR: the hazard is exactly as specified here (i.e., endorsing a 'fact', albeit a different one) and thus links to the same type of hazardous event and the same type of harm. It is not clear that the risk involved in remaining silent is Hempel's IR, since the hazard 'remaining silent' is different and cannot link to the same hazardous event (i.e., 'remaining silent' cannot link to 'endorsing a 'fact' whose objective truth value is false). Regardless, what *is* clear is that when deciding whether or not to endorse a fact, one must weigh the relative harms and benefits of endorsing it, endorsing its negation, or remaining silent. We think the significance of this ultimately loads into the hazard of Hempel's IR as we define it here.
9. Just for example: "Jeffrey proposed that scientists should assign probabilities to hypotheses in light of the available evidence and pass these probabilities along to policy makers." (Steel 2015, 81)"; "Jeffrey asserted that probabilities in a Bayesian approach are not the sort of thing one chooses to accept or reject; *they are degrees of belief scientists have and which they should report to policy makers*." (Steel 2015, 81, italics in the original)

10. Note, inter alia, that here Jeffrey is tipping his hand that, according to him, like Hempel, even singular statements like "it is raining outside" are subject to IR if accepted or rejected. It is just lawlike hypotheses, that are "least of all" immune to it.
11. Although on some rationalist accounts, there will be no risk at all, because determining the probability will be a matter of a uniquely correct logic, as we acknowledge below.
12. Ramsey and de Finetti are both clear in various places that this follows from their view. See Shafer (1981) and Gibbard (2007), especially section 3, for more recent discussions. Decision theory on the part of the expert will quickly become a matter of game theory between expert and customer, because the customer will soon realize that probabilities are being given strategically, and hijinx will ensue. How to deal with this when eliciting credences from experts was first discussed by Brier (1950).
13. Following Rausand (2011), a hazard is "a source of danger that may cause harm to an asset", a hazardous event "the first event in a sequence of events that, if not controlled, will lead to undesired consequences (harm)" (598-99). As noted, harm can be understood broadly as undesired consequences.
14. See e.g., Rausand (2011): "external events and conditions may influence the event sequences" (129).
15. This point relates to the well-known issue of 'explanatory' versus 'pragmatic' clinical trials, also discussed by Bluhm (2017), in which participant eligibility criteria are more or less strict by design.
16. To confirm, Douglas (2009, 96) herself uses both the terms 'claim' and 'outcome' in defining the direct/indirect role distinction: e.g., "values can act as reasons in themselves to accept a claim, providing direct motivation for the adoption of a theory"; "In the direct role, values determine our decisions in and of themselves, acting as stand-alone reasons to motivate our choices. They do this by placing value on some intended option or outcome, whether it is to valorize the choice or condemn it."

## WORKS CITED

Alexandrova, Anna. 2010. "Adequacy for Purpose: The Best Deal a Model Can Get." *The Modern Schoolman* 87 (3/4): 295–301.

Anderson, James A. 2006. "The Ethics and Science of Placebo-Controlled Trials: Assay Sensitivity and the Duhem–Quine Thesis." *The Journal of Medicine and Philosophy* 31 (1): 65–81.

Bluhm, Robyn. 2017. "Inductive Risk and the Role of Values in Clinical Trials." In *Exploring Inductive Risk: Case Studies in Values in Science*, edited by Kevin C. Elliott and Ted Richards, 171–91. New York: Oxford University Press.

Bokulich, Alisa, and Wendy S. Parker. 2021. "Data Models, Representation, and Adequacy-for- Purpose." *European Journal for Philosophy of Science* 11: 31.

Biddle, Justin, and Rebecca Kukla. 2017. "The Geography of Epistemic Risk." In Exploring Inductive Risk: Case Studies in Values in Science, edited by Kevin C. Elliott and Ted Richards, 216–37. New York: Oxford University Press.

Biddle, Justin. 2018. "Antiscience Zealotry"? Values, Epistemic Risk, and the GMO Debate." *Philosophy of Science* 85 (3): 360–79.

Biddle, Justin. 2020a. "On Predicting Recidivism: Epistemic Risk, Tradeoffs, and Values in Machine Learning." *Canadian Journal of Philosophy* 1–21.

———2020b. "Epistemic Risks in Cancer Screening: Implications for Ethics and Policy". *Studies in History and Philosophy of Biological and Biomedical Sciences* 79: 101200.

Biddle, Justin, and Eric Winsberg. 2009. "Value Judgments and the Estimation of Uncertainty in Climate Modeling." In *New Waves in the Philosophy of Science*, edited by P. D. Magnus and Jacob Busch, 172–197. New York: Palgrave MacMillan.

Brier, Glenn W. 1950. "Verification of Weather Forecasts Expressed in Terms of Probability." *Monthly Weather Review* 78 (1) :1–3. https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2

Brown, Matthew J. 2017. "Values in Science: Against Epistemic Priority." In *Current Controversies in Values and Science*, edited by Kevin Elliott and Daniel Steel, 64–78. London: Routledge.

Bueter, Anke. 2015. "The irreducibility of value-freedom to theory assessment." *Studies in History and Philosophy of Science* 49: 18–26.

de Melo-Martín, Inmaculada, and Kristen Intemann. 2016. "The Risk of Using Inductive Risk to Challenge the Value-Free Ideal." *Philosophy of Science* 83 (4): 500–20.

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: Humphreys, P. (Ed.), The Oxford Handbook of Philosophy of Science, edited by Paul Humphreys, 609–630. New York: Oxford University Press.

Drummond, Michael F., Mark J. Sculpher, Karl Claxton, Greg L. Stoddart, and George W. Torrance. 2015. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press.

Elliott, Kevin C. 2013. "Douglas on Values: From Indirect Roles to Multiple Goals." *Studies in History and Philosophy of Science Part A* 44 (3): 375–83.

Elliott, Kevin C., and Daniel J. McKaughan. 2009. "How Values in Scientific Discovery and Pursuit Alter Theory Appraisal." *Philosophy of Science* 76 (5): 598–611.

Elliott, Kevin C., and Ted Richards. 2017a. *Exploring Inductive Risk: Case Studies of Values in Science*. Oxford University Press.

———. 2017b. "Exploring Inductive Risk: Future Questions" In *Exploring Inductive Risk: Case Studies in Values in Science*, edited by Kevin Elliott and Ted Richards, 261–77. New York: Oxford University Press.

Fricker, Miranda. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. New York: Oxford.

Frigg, Roman, and Stephan Hartmann. 2020. "Models in Science." In *Stanford Encyclopedia of Philosophy,* edited by Edward N. Zalta. https://plato.stanford.edu/entries/models-science/

Galavotti, Maria Carla. 1991. "The Notion of Subjective Probability in the Work of Ramsey and de Finetti." *Theoria* 57 (3): 239–59. https://doi.org/10.1111/j.1755-2567.1991.tb00839.x.

Gastala, Nicole M., Peter Wingrove, Anne Gaglioti, Stephen Petterson, and Andrew Bazemore. 2016. "Medicare Part D: Patients Bear The Cost Of 'Me Too' Brand-Name Drugs." *Health Affairs* 35 (7): 1237–40.

Gibbard, Allan. 2007. "Rational Credence and the Value of Truth." *Oxford Studies in Epistemology* 2: 143–64.

Glasser, Stephen P., and George Howard. 2006. "Clinical Trial Design Issues: At Least 10 Things You Should Look for in Clinical Trials." *Journal of Clinical Pharmacology* 46 (10): 1106–15.

Harvard, Stephanie, Gregory R. Werker, and Diego. Silva. 2020. "Social, Ethical, and Other Value Judgments in Health Economics Modelling." *Social Science and Medicine* 253: 1–9.

Hempel, Carl G. 1935a. "On the Logical Positivists' Theory of Truth." *Analysis* 2 (4): 49–59.

———. 1935b. "Some Remarks on 'Facts' and Propositions." *Analysis* 2 (6): 93–6.

———. 1954. "A Logical Appraisal of Operationism." *The Scientific Monthly* 79 (4): 215–20.

———. 1965. "Science and Human Values." In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, 81–96. New York: The Free Press.

———. 1981. "Turns in the Evolution of the Problem of Induction." *Synthese* 46 (3): 389–404.

———. 2000. "The Problem of Truth"." In *Selected Philosophical Essays*, edited by Carl Gustav Hempel and Richard Jeffrey, 35–74. New York: Cambridge University Press.

Intemann, Kristen. 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.

Kidd, Ian, Gaile Pohlhaus, and Jose Medina. 2017. *The Routledge Handbook of Epistemic Injustice*. New York: Routledge.

Leitgeb, Hannes, and André Carus. 2020. "Supplement to Rudolph Carnap B. Carnap versus Quine on the Analytic-Synthetic Distinction." In *Stanford Encyclopedia of Philosophy,* edited by Edward N. Zalta. https://plato.stanford.edu/entries/carnap/carnap-quine.html

Longino, Helen E. 1990. *Science as Social Knowledge: Values and Objectivity in Scientific*
*Inquiry.* Princeton: Princeton University Press.

Light, Donald W., and Joel R. Lexchin. 2020. "Pharmaceuticals as a Market for 'Lemons': Theory and Practice." *Social Science and Medicine* 268: 113368. https://doi.org/10.1016/j.socscimed.2020.113368.

Lundh, Andreas, Joel Lexchin, Barbara Mintzes, Jeppe B. Schroll, and Lisa Bero. 2017. "Industry Sponsorship and Research Outcome." *Cochrane Database of Systematic Reviews* 2: : MR000033. https://doi.org/10.1002/14651858.MR000033.pub3.

Mann, Howard, and Benjamin Djulbegovic. 2013. "Comparator Bias: Why Comparisons Must Address Genuine Uncertainties." *Journal of the Royal Society of Medicine* 106 (1): 30–3. https://doi.org/10.1177/0141076812474779.

Mosor, Erika, Michaela Stoffer-Marx, Günter Steiner, et al. 2020. "I Would Never Take Preventive Medication! Perspectives and Information Needs of People Who Underwent Predictive Tests for Rheumatoid Arthritis." *Arthritis Care and Research* 72 (3): 360–8. https://doi.org/10.1002/acr.23841.

Parker, Wendy S. 2010. "Whose Probabilities? Predicting Climate Change with Ensembles of Models." *Philosophy of Science* 77 (5): 985–97.

———. 2020. "Model Evaluation: An Adequacy-for-Purpose View." *Philosophy of Science* 87 (3): 457–77.

Parker, Wendy S., and Greg Lusk. 2019. "Incorporating user values into climate services." *Bulletin of the American Meteorological Society* 100 (9): 1643–50.

Parker, Wendy S., and Eric Winsberg. 2018. "Values and Evidence: How Models Make a Difference." *European Journal for Philosophy of Science* 8 (1): 125–42.

Peschard, Isabelle F., and Bas C. van Fraassen. 2014. "Making the Abstract Concrete: The Role of Norms and Values in Experimental Modeling." *Studies in History and Philosophy of Science Part A* 46: 3–10.

Powers, Jack. 2017. "The Inductive Risk of 'Demasculinization'." In Exploring Inductive Risk: Case Studies in Values in Science, edited by Kevin Elliott and Ted Richards, 239–60. New York: Oxford University Press.

Quine, Willard V. 1951. "On Carnap's Views on Ontology." *Philosophical Studies* 2 (5): 65–72.

Rausand, Marvin. 2011. *Risk Assessment: Theory, Methods, and Applications*. New Jersey: John Wiley and Sons.

Rudner, Richard. 1953. "The Scientist qua Scientist Makes Value Judgments." *Philosophy of Science* 20 (1): 1–6.

Shafer, Glenn. 1981. "Constructive Probability." *Synthese* 48: 1–60.

Stanev, Roger. 2017. "Inductive Risk and Values in Composite Outcome Measures." In *Exploring Inductive Risk: Case Studies in Values in Science*, edited by Kevin Elliott and Ted Richards, 171–91. New York: Oxford University Press.

Steel, Daniel. 2017. "Qualified Epistemic Priority: Comparing Two Approaches to Values in Science." In *Current Controversies in Values in Science*, edited by Kevin Elliott and Daniel Steel, 49–63. New York: Routledge.

Steel, Daniel. 2015. "Acceptance, values and probability." *Studies in History and Philosophy of Science* 53: 81–8.

Steel, Daniel, and Kyle Powys Whyte. 2012. "Environmental Justice, Values, and Scientific Expertise." Kennedy Institute of Ethics Journal 22 (2): 163–82.

Valles, Sean A. 2018. *Philosophy of Population Health: Philosophy for a New Public Health Era*. New York: Routledge.

Van Fraassen, Bas C. 1980. *The Scientific Image*. Oxford: Oxford University Press.

———. 2008. *Scientific Representation: Paradoxes of Perspective*. Oxford: Oxford University Press.

Wilholt, Torsten. 2009. "Bias and Values in Scientific Research." *Studies in History and Philosophy of Science Part A* 40 (1): 92–101.

Winsberg, Eric. 2010. *Science in the Age of Computer Simulation*. Chicago: University of Chicago Press.

———. 2012. "Values and Uncertainties in the Predictions of Global Climate Models." *Kennedy Institute of Ethics Journal* 22 (2): 111–37.

———. 2018. *Philosophy and Climate Science*. Cambridge: Cambridge University Press.

Winsberg, Eric, Bryce Huebner, and Rebecca Kukla. 2014. "Accountability and Values in Radically Collaborative Research." *Studies in History and Philosophy of Science Part A* 46: 16–23.

Winsberg, Eric, Naomi Oreskes, and Elisabeth Lloyd. 2020. "Severe Weather Event Attribution: Why Values Won't Go Away." *Studies in History and Philosophy of Science Part A* 84: 142–49. https://doi.org/10.1016/j.shpsa.2020.09.003.