Values in Science: What Are Values Anyway?

Kevin C. Elliott* and Rebecca Korf†

*Lyman Briggs College, Dept. of Fisheries and Wildlife, and Dept. of Philosophy

Michigan State University

kce@msu.edu

†Department of Logic and Philosophy of Science

University of California, Irvine

rkorf@uci.edu

Abstract: Although the philosophical literature on science and values has flourished in recent years, the central concept of "values" has remained ambiguous. This paper endeavors to clarify the nature of values as they are discussed in this literature and then highlights some of the major implications of this clarification. First, it elucidates four major concepts of values and discusses some of their strengths and weaknesses. Second, it clarifies the relationships between these concepts of values and a wide variety of related concepts that are sometimes used interchangeably in the philosophical literature. Third, it argues that this conceptual clarification reveals that much of the literature on science and values has discussed different concepts of values without making these differences clear. The paper illustrates this point by analyzing the different concepts of values at play in different arguments against the value-free ideal and in proposals for managing values. Understanding the literature on values in science as a patchwork of related discourses rather than a single discourse can help researchers more thoughtfully choose a concept of values that best fits their philosophical targets and goals, rather than conflating different discourses because of the common terminology of "values."

Values in Science: What Are Values, Anyway?

Kevin C. Elliott and Rebecca Korf

1. Introduction

Although the philosophical literature on science and values has flourished in recent years, the central concept of "values" has remained ambiguous.[1] As one of us previously noted in an overview of the literature on this topic, "those writing about values in science often lament that the word 'value' is used as a label for a very wide array of phenomena that ought to be treated in different ways" (Elliott 2022, 4). Justin Biddle provides a particularly pointed statement of this worry:

> [T]here are a wide range of different factors that can fill the gap between "insight" (i.e., logic, evidence, and epistemic values broadly construed) and decision making in science. These factors might be sociological in nature (e.g., in one particular scientific sub-discipline, one set of norms are typically employed for evaluating research, as opposed to some other set, which are employed in some other sub-disciplines); they might be consciously adopted ethical or political values; they might be unconsciously held subjective preferences or ideological assumptions, and so on. Some of these factors, again, are properly described as values, while others are not. One of the tasks for philosophers of science working in this area should be to develop a more fine-grained understanding of the kinds of factors that can operate here, so as to specify more clearly which factors play a legitimate role in science and which do not…. (Biddle 2013, 131-132)

Of course, those working on this topic have not left the concept of values completely ambiguous and undefined. Kuhn (1977) defined values as criteria for theory choice. He labeled these criteria "values" because he said that, in contrast to "rules," they merely *influence* theory choice rather than forcing or *determining* it (1977, 111). McMullin (1983) provided a particularly extensive clarification of value concepts in his Presidential Address for the Philosophy of Science Association. He distinguished "emotive" concepts of value from those that treat values as desirable characteristics (e.g., Kuhn's concept of values as criteria for theory choice). He also noted that values could play the role of ideals (e.g., truth as an ideal of science), ethical standards, and utilities (as in decision theory). Finally, he acknowledged that "externalist historians" sometimes also placed "not only social and personal goals but also elements of world-view, metaphysical, theological, and the like" under the umbrella of "values" (1983, 19).

---

[1] Throughout this paper, we will use the terms "concept" and "conception" in much the same way that political philosophers often use these terms. They commonly distinguish a concept, which is a general idea, from a conception, which is a further specification of that concept (see e.g., Swift 2001). Thus, we will refer to four main "concepts" of value throughout the paper, and then we will talk about particular philosophers' individual conceptions of those main concepts. We thank Ahmad Elabbar for helping us to clarify this point.

However, McMullin resisted this move if those elements were treated as purely psychological or social causes rather than as cognitive criteria for theory choice.

Despite these initial efforts to define values, as well as a range of subsequent efforts (see e.g., Brown 2020; Elliott 2017; Lacey 1999), the nature of values has remained elusive in practice. As McMullin himself acknowledged, this is partly because "'Value' is one of those weasel words that slip in and out of the nets of the philosopher" (1983, 4). Even though there have been some efforts to define the concept of "value" in a relatively narrow fashion, Helen Zhao notes that "It is an open secret that 'value' has been promiscuously employed by philosophers of science" (2022, 3). According to Miriam Solomon:

> "Value" has been used to include political values, aesthetic preferences, psychological biases, cognitive goals, personal and societal goals, ideologies, and pre-theoretic intuitions. So "value" is, in practice, not restricted to ethical values or even aesthetic values. "Values" include pre-theoretic assumptions, ethical conduct of inquiry, and causes of preference for one theory over another. (2012, 332-333)

Philosophers of science working in this area have recently begun trying to bring greater clarity to discussions of values. For example, in her efforts to distinguish the different senses in which scientific judgments can be "value-laden," Zina Ward (2021) clarified that values can serve as either motivating reasons or as justifying reasons, and values can also serve as either causal effectors or as affected goods. Similarly, Helen Zhao (2022) has distinguished two different ways of conceptualizing values. First, she claims that values are sometimes characterized as factors that play a *causal role* in scientific decision making; she calls this the "Decision Vector" account of values. Second, she suggests that values are sometimes regarded as aims that scientists should have; she calls this the "Goal" account of values. Eric Winsberg (2024) has also recently attempted to clarify the nature of values in science by treating all values as preferences over outcomes (i.e., utilities in decision theory) and denying that what are traditionally called "epistemic values" actually count as values. Finally, Hannah Hilligardt (2022) recently drew on the work of Iris Marion Young in an effort to distinguish between values and three related phenomena: social perspectives, opinions, and interests.

We aim to build on this recent work in order to develop a more systematic account of the nature of values in science and their relationships with other phenomena. In Section 2, we clarify four different ways of conceptualizing values. Rather than affirming one particular concept as the "right" way to think about values, we discuss the strengths and weaknesses of the four approaches in an effort to foster more productive discussions about values in the future. Section 3 then highlights the wide array of other concepts that are related to values and sometimes treated as synonymous with values. By discussing these concepts, we aim to advance Biddle's goal of clarifying the array of factors that can play a role in scientific reasoning and reflecting on how those factors should be characterized. Finally, Section 4 explores the broader significance of this conceptual clarification for the literature on science and values. The section argues that attention to different concepts of values highlights ways in which the literature on science and values actually consists of a patchwork of discourses that address distinct but related phenomena.

Although we think it is fruitful to distinguish between the different concepts and phenomena that fall under the broad umbrella of "science and values," we are not calling for the literature to fracture along these conceptual lines. We think it is useful for those working in this research area to continue employing a variety of different concepts in order to capture the diversity of phenomena that they study. Nevertheless, in order to avoid confusion and move the discussion forward in a coherent fashion, it is crucial for those studying this topic to clarify which concepts of values are at play in their work.

2. Major Concepts of Values

In an effort to bring greater clarity to the confusing proliferation of value concepts, our goal in this section is to elucidate four major ways of conceptualizing values. We should emphasize that we are focusing on the way the term 'value' is used in the literature on science and values; we do not presume to provide an exhaustive analysis of the way 'value' is used across all areas of inquiry. The boundaries between these four concepts are not meant to be sharp distinctions; while we believe these four categories capture major themes in the literature, there is often considerable overlap between them, as we will discuss later in this section. In addition, we will not attempt to decide which concept is best; rather, we will attempt to clarify the major options available in the literature so that those working on the topic of science and values can be more careful in their choice and specification of their preferred concept. As we will see in Section 4, the failure to provide this specification risks running together distinct discourses and causing significant confusion.

Values as Criteria or Standards for Choice

First, one of the most common ways to conceptualize values is in terms of *criteria or standards for choice* (see Table 1). We have already seen that Kuhn (1977) characterized values this way in his classic essay that helped to launch discussions of science and values. He identified five characteristics of good scientific theories: accuracy, consistency, broad scope, simplicity, and fruitfulness. Kuhn claimed that these properties of scientific theories are all "standard criteria for evaluating the adequacy of a theory" (1977, 103). But he went on to say that the criteria operate as values insofar as different scientists can differ in assessing the extent to which the criteria are met and their relative importance. McMullin followed Kuhn in focusing primarily on values in science as criteria for theory choice. Other philosophers of science have expressed similar views that characterized values as criteria for evaluating a broader array of entities; for example Hugh Lacey claimed that "Values are properties of an object that are deemed to be criteria for appraising its value…" (2017, 15). In her book *Value in Ethics and Economics*, Elizabeth Anderson (1993) also referred to values as criteria in this wider sense: "Call a person's values whatever standards she accepts for evaluating persons, actions and things" (1993, 3).

Table 1. Four ways of conceptualizing values, with major examples of each kind.

Criteria or standards for choice	Causal factors	Beliefs or attitudes about desirable things	Desirable things
Kuhn (1977) McMullin (1983) Lacey (1999) Anderson (1993)	Longino (1990) Solomon (2012) Zhao (2022; 2023) McMullin (1983)	Steel (2010) Dietz (2013) Karwat (1982) Brown (2020) Winsberg (2024)	Vaske and Manfredo (2012) Elliott (2017) Brown (2020)

Values as Causal Factors

But not everyone working in this area talks about values in this way. As noted in the preceding quotation from Solomon (2012, 332), some people use the term 'value' to refer to "political values, aesthetic preferences, psychological biases, cognitive goals, personal and societal goals," and so on. These "values" do not all operate as criteria for choice. Rather, they operate as *factors that causally influence scientists' choices, habits, and practices* (see e.g., Zhao 2023).[2] This is a second way to conceptualize values in science. Helen Zhao (2022) traces the expansion of the concept of "value" from criteria to causal factors back to McMullin's (1983) Presidential Address. As we have seen, McMullin wanted to focus specifically on criteria for theory choice, but he noted that the concept of "value" could also encompass political, social, and religious values, as well as other idiosyncratic influences on scientists. McMullin separated these other "non-epistemic" values from the "epistemic" values on which he wanted to focus, but Zhao contends that he opened the door to thinking about values in a broader way. Therefore, she proposes a concept of values as "decision vectors," i.e., as "factors that play some causal role in scientific decision-making" (2022, 6; see also Zhao 2023).[3] For example, she notes that scientists could pursue a theory because it is predictively accurate and/or because it promotes their job security. Both act as values in the sense that they causally influence the scientists' choices. Others who take this wider view of values include Helen Longino, who argues that the background assumptions that impact scientific decisions can "encode" values (1990, 128), where this encoding relationship can involve values causing scientists to accept the assumptions even when the values are not consciously recognized.

Values as Beliefs or Attitudes about Desirable Things

---

[2] This definition is very broad, so to make this concept of values compelling it might be necessary to narrow the range of causal factors that actually count as values. For example, one could limit this concept so it focuses only on *psychological* causal factors and not on causal factors "external" to scientists' mental processes. However, for the purposes of this paper, we will leave this description of values broad (as those who employ this concept of values typically do) so we do not beg the question of what factors do or do not count as values.

[3] Zhao borrows decision vectors as a "term of art" from Miriam Solomon (2001). While Solomon develops several possible models that specify the magnitude and direction of decision vectors, Zhao uses the term in a general sense.

Some would argue, however, that the concept of values as causal factors is too broad. This seems to be Biddle's (2013) view when he argues that not all factors that fill the gap between evidence and scientific conclusions should be labeled values. One way to narrow this concept of values is to focus not on all causal factors that influence scientists but only on a subset of them, namely, those that constitute *beliefs or attitudes about what is desirable* (see Steel 2010). This seems to be the way many social scientists think about values. For example, sociologist Thomas Dietz quotes a classic psychological definition of values as "(a) concepts or beliefs, (b) about desirable end states or behaviors, (c) that transcend specific situations, (d) guide selection or evaluation of behavior and events, and (e) are ordered by relative importance" (2013, 1). Similarly, in an article for political scientists, Miroslaw Karwat characterizes values as "ideas of needs," i.e., expressions of a "subject's attitude toward its own needs" (1982, 198).

Matthew Brown may express something like this view in his book *Science and Moral Imagination* (2020). He grounds values in the activity of valuing, which involves being "disposed to act" for something (2020, 115). According to Brown, that disposition to act expresses a desire (i.e., a value), but he notes that the desire can take different forms, ranging from a conscious desire for something to a more "deeply ingrained" attitude to a more idealistic "desire-to-desire" things that are not yet consciously desired (2020, 116). In a recent article, Eric Winsberg adopts another approach that seems to fit within this third category. He argues that values "... *are desired outcomes.* More precisely, they are the numerically weighted *preferences* decision makers (including scientists) have *over prospects*" (2024, 6, italics in original). In other words, he conceptualizes scientists as decision makers operating within the framework of decision theory. According to that framework, decision makers act on the basis of two kinds of information: (1) factual information about possible outcomes and their likelihood; and (2) values or utilities that express how much those outcomes are desired. For Winsberg, values in science should simply be regarded as the utilities that express scientists' preferences regarding various outcomes.

Values as Desirable Things

But these three concepts of values still do not capture all the major ways in which they have been discussed in the literature; a fourth approach is to characterize values not as beliefs or ideas about desirable things but as the desirable things themselves. Some social scientists appear to hold this view. For example, Vaske and Manfredo state that "Values are commonly defined as desirable individual end states, modes of conduct, or qualities of life that we individually or collectively hold dear, such as freedom, equality, and honesty" (2012, 43). Similarly, Kevin Elliott defines values as "things that are desirable or worthy of pursuit" (2017, 11). He intended this definition to cover both criteria (properties of theories that are desirable) as well as desirable states of affairs (pers. comm.). Brown (2020) may be somewhat similar to Elliott in trying to capture desirable things in themselves as values along with other kinds of entities. For example, Brown's list of values in his book includes criteria like consistency and systematicity as well as desirable phenomena like peer review and providing due credit for

scientific work (2020, 127-129). For those who view values as goals, those goals could also be interpreted as the things themselves that are being pursued (see Zhao 2022).

Relationships between the Concepts

Reflecting on these four concepts of values, one might notice that they have similarities to Ward's (2021) clarification of four ways in which choices in science can be value-laden. For example, the concept of values as criteria or standards is similar to Ward's (2021) notion that values can act as "justifying reasons." The concept of values as causal factors appears to align with Ward's (2021) notion that values can serve as causal effectors. The concept of values as beliefs or attitudes about what is desirable has some similarities to Ward's idea that values can serve as motivating reasons. And the concept of values as desirable things accords very well with Ward's notion that values can serve as affected goods. However, we think it is important to distinguish concepts of what values *are* from accounts of what they can *do* or the *roles they can play*. For example, values that fall under a single concept (e.g., a belief or attitude about what is desirable) could play different roles (e.g., serving as a causal effector, a justifying reason, or a motivating reason) in different contexts. Similarly, a single role for values (e.g., serving as a causal effector) could be played by values falling under different concepts (e.g., a standard, a causal factor, or a belief/attitude about what is desirable). Thus, although it is fruitful to recognize that there are significant similarities between the major concepts of values that we have identified and the major roles or effects for values that Ward identified, we think it is important not to simply equate what values *are* with what they can *do*.

It is also noteworthy that there are many connections between our four concepts of values and that the boundaries between them are somewhat indistinct. For example, we regard criteria (our first concept of values) as *reasons for choice*. This concept has close relationships with our fourth concept of values (i.e., desirable things) because the reasons for choice often refer to properties of theories or hypotheses that are desirable. In addition, the concept of values as criteria for choice is closely related to our third concept of values (beliefs or attitudes about what is desirable), insofar as people often have beliefs about the desirable properties that should serve as criteria for choice. There are also close relationships between our second and third concepts of values because beliefs or attitudes about desirable qualities often act as causal factors that influence the decisions of scientists.

Despite these close relationships, we think it is still important to recognize distinctions among these concepts. For example, although criteria for choice can be manifested in the form of beliefs about desirable qualities, they do not have to be. Institutional systems can be set up to establish criteria for choice even if individual scientists do not believe that what they promote is desirable. Similarly, beliefs about desirable qualities do not always act as criteria, or even as causal factors, and there can be many kinds of causal factors that would not be classified as beliefs or attitudes. For example, whereas beliefs or attitudes are typically regarded as relatively stable commitments that we would endorse upon reflection, other causal factors need not involve the same sort of commitment. Although one could put a great deal of effort into providing precise definitions for each of our four concepts and delineating the precise relationships

between them, that is not our project here. We think that different figures are likely to define each concept slightly differently, so at this point we think it is more fruitful to draw attention to these general concepts of values and the different roles they tend to play in the values-and-science literature rather than worrying about the exact relationships among them. Throughout the paper, we will see that even though the four concepts sometimes overlap, different concepts receive pride of place in different parts of the literature on science and values, and they each have their own strengths and weaknesses.[4]

Assessing the Concepts

Consider, for example, the strengths of thinking about values as criteria or standards. This has the advantage of connecting with the classic work of Kuhn and McMullin, who clarified that scientific reasoning involved appeals to desirable characteristics that could not easily be weighed in a rule-governed fashion. However, a disadvantage of this approach is that it may not adequately capture the full range of ways that values are currently discussed in the values and science literature. For example, when Heather Douglas (2009; 2022) contends that scientists should consider ethical values when deciding how to respond to inductive risk, she does not seem to regard those values as criteria for theory choice in the same way that Kuhn and McMullin saw them. They saw these criteria as properties of theories themselves, whereas Douglas's ethical values involve beliefs about how to handle the social consequences of accepting or rejecting theories. Thus, it seems more appropriate to associate Douglas's notion of values with our third concept (i.e., beliefs or attitudes about what is desirable).[5] Similarly, when Helen Longino says that background beliefs can "express," "encode," or "be motivated by" background beliefs (1990, 86, 128), she does not appear to be referring to values as criteria. Rather, she appears to be thinking about values as causal factors that contribute to accepting background beliefs that go on to justify scientific judgments.

A strength of conceptualizing values as causal factors is the breadth of this concept. This breadth is especially helpful because it captures the way scientists sometimes seem to think about values as any factors other than evidence that have an influence on scientific reasoning or practices. As mentioned above, McMullin (1983) suggested that some historians have also thought about values in this very broad way, and some philosophers may also take something like this view (see e.g., Hudson 2021). Another strength of this concept is that it

---

[4] It is also worth noting that different concepts have received different emphasis at particular historical moments. For example, we have seen that the early work of Kuhn and McMullin in the 1970s and 1980s emphasized values as criteria or standards, whereas Longino's work in the 1990s emphasized values as causal factors. In the 2000s and 2010s, concepts of values as beliefs or attitudes about what is desirable and desirable things in themselves gained prominence. All four concepts can be found in the current literature.

[5] Admittedly, inductive risk is often handled in practice by setting standards or guidelines for what counts as sufficient evidence to accept a conclusion. These standards or guidelines could be regarded as criteria for choice, and thus they could fall under our first concept of values. However, a primary emphasis in Douglas's work is that scientists have ethical responsibilities to assess the appropriateness of these criteria, and this assessment process typically makes reference to underlying beliefs or attitudes about what is desirable. Thus, we contend that beliefs or attitudes about what is desirable are central to Douglas's concept of values. We thank Zina Ward for helping us to think more carefully about this issue.

helps to explain one of the ways in which background beliefs or assumptions can be "value-laden" in the sense described by Longino (1990)—namely, those background beliefs could be caused by particular psychological or social factors. However, as noted above, the weakness of this concept is that it lumps together a wide variety of different phenomena that should perhaps be analyzed differently (see Biddle 2013). For example, Douglas and Elliott (2022) have argued that biases and values should be distinguished, but this distinction could easily be lost if all causal influences on scientists' reasoning are conceptualized together as values.

As noted above, one way to develop a narrower concept of values is to focus on people's beliefs or attitudes about what is desirable. In addition to narrowing the scope of things that count as values, this approach has the virtue of connecting with the ways many social scientists talk about values. In addition, this concept helps to clarify the nature of values in the "aims" approach to values in science (see e.g., Elliott 2013; Intemann 2015; Hicks 2022). The aims approach has historically been somewhat ambiguous about the nature of aims or goals and their relationship to values. If one adopts the concept of values as people's beliefs about what is desirable, then aims could count as values under this third concept, while the criteria for selecting hypotheses, theories or models that help to achieve those aims could count as values under the first concept considered in this section.

Nevertheless, conceptualizing values as beliefs or attitudes about what is desirable may also face weaknesses. One of these weaknesses is that in order to characterize values or value-ladenness, this concept depends directly on individuals and their perspectives on what is desirable. This could be a disadvantage because these beliefs and attitudes can be difficult to discern, whether by individuals themselves or by scholars trying to analyze their decisions.[6] Another challenge arises as the literature on science and values begins to turn its attention more directly to scientific practices and institutions rather than individual scientists and their beliefs and choices (see e.g., Douglas 2018; Zhao 2023). To capture the ways in which these practices and institutions are value-laden, it may be necessary to highlight their effects on desirable things (thereby appealing to the fourth concept of values) rather than focusing solely on specific people's beliefs or attitudes. However, it is doubtful that a fully adequate account of values in science could be developed that focused solely on desirable things without ever referring to scientists' beliefs about them

One of the fundamental lessons of this paper is that philosophers should be intentional about choosing the concept of "values" that best fits their goals. Given the strengths and weaknesses that we have identified in this section, some concepts of values are better suited for certain aims and contexts than others. For example, when one is trying to characterize the full array of factors that could influence a scientific debate (such as in some of Solomon's (2001) work), conceptualizing values as causal factors could be particularly fruitful. In contrast, when one is advising individual scientists about how to reason ethically about their choices (e.g.,

---

[6] One might object that values of this sort can still exist and influence scientists' reasoning even if they are difficult to consciously identify. However, if these values are difficult to identify or isolate, it might be more fruitful to analyze scientific decisions in terms of their effects on desirable things (i.e., our fourth concept of values) rather than based on more elusive beliefs or attitudes.

Brown 2020; Douglas 2021), it makes much more sense to conceptualize values as beliefs about what is desirable. In policy contexts, it might be more helpful to attempt to characterize the goals at play (treating those goals as values in the sense of desirable things) and clarify the criteria best suited for evaluating scientific practices in order to achieve those goals (treating values as criteria).

Thus, the varied concepts of values in the values-and-science literature are probably warranted given the variety of tasks that philosophers of science aim to achieve. And even if one were to conclude that one of these concepts was somewhat problematic, we would caution against being too quick to abandon it completely or treat it as irrelevant for the literature on values in science. For example, even if one were to conclude that some causal factors should not be treated as values in the strict sense of the word, those factors might still merit reflection regarding the roles they should play in scientific practice and the ways they can be most fruitfully managed. Nevertheless, despite the benefits of being able to work with multiple concepts of values, we will argue in Section 4 that achieving greater clarity about which concept of values is at play is often important, especially when formulating arguments about the proper roles for values in science and the best strategies for managing values. Before delving into this discussion of the philosophical implications of our argument, however, we aim to provide some additional conceptual clarification by elucidating the relationships between values and other concepts that are sometimes related to them.

3. Values and Other Concepts

Although we think the distinctions provided in Section 2 between different concepts of values help to provide a great deal of clarity, these distinctions still do not fully explain the relationships between values and a variety of other related concepts. Consider the following: ideals, opinions, subjective factors, utilities, epistemic criteria, standards, guidelines, positionality, interests, preferences, biases, virtues, worldviews, identities, ideologies, aims, goals, professional culture, organizational climate, purposes, norms, and background beliefs/assumptions. Because the concept of values has been used in such broad ways in some previous literature, it is somewhat unclear whether these other phenomena count as values, and there can be costs to this lack of clarity.[7] As Ward puts it:

> Much philosophical work on science and society has been framed in terms of values, perhaps because we lack the conceptual resources to do otherwise, or perhaps because "values in science" is a well-established topic in the field. This trend has unfortunate consequences. Shoehorning everything related to science and the social into the

---

[7] We are focusing here on concepts that could plausibly be equated with or treated synonymously with values. One could extend our list even further if one included all the phenomena that can be "value-laden," such as structures, laws, policies, institutions, and practices. These are very important phenomena to investigate, but our focus in this paper is on concepts that someone might actually equate with values.

literature on values in science has distorted the targets of investigation and hidden the multiplicity of ways in which choices can involve values. (Ward 2021, 61)[8]

With these concerns in mind, this section aims to clarify the conceptual landscape and provide some greater precision about how other concepts relate to the four concepts of values discussed in Section 2.

There are at least four reasons that the relationships between values and other concepts have been a source of confusion. First, as discussed in Section 2, there are multiple different concepts of values under which other concepts could fall. Second, other concepts do not always fall under a single concept of values. As we will discuss below, some concepts (e.g., worldviews) could act as causal factors (or incorporate elements that act as causal factors) in some cases while serving primarily as beliefs about what is desirable in other cases. Third, as we will explain further, a concept could refer to phenomena that fall under one of the concepts of values in some cases but that do not count as a value at all in other cases. Finally, as noted above, not everyone would agree that all entities falling under one or more of the four concepts from Section 2 should be counted as values, so even if a concept is associated with one of those concepts, there could still be disagreement about whether it actually counts as a value.[9]

Given these complications, attempting to clearly delineate whether specific phenomena count as values is likely to prove difficult and go beyond the scope of this paper. What does seem feasible is to clarify whether these phenomena could, in at least some cases, fall under a particular concept from Section 2 (see Table 2). By providing this sort of analysis, we can develop a better understanding of how these phenomena relate to values and why one might or might not want to characterize them as values.

Table 2. Tentative categorization of concepts related to values

Criteria or standards for choice	Factors that play causal roles	Beliefs or attitudes about desirable things	Desirable things	Other
Standards Guidelines Virtues	Biases Emotions Interests	Organizational climate Professional culture	Goals Aims Purposes	Background beliefs or assumptions

---

[8] We thank Ahmad Elabbar for providing a potential example of how debates could be impoverished by "shoehorning" all social factors under the concept of values. He notes that in a situation where a company has an interest in polluting a river and harming the health of residents along the river, it may be helpful to distinguish the "interests" of the company from the "values" of the citizenry. One might think that interests and values should be treated differently by regulators rather than lumping them all together as values to be weighed against each other.

[9] To add an additional layer of confusion, some of the work situated firmly in the values-in-science literature employs other concepts listed here rather than the concept of values, even though the phenomena described plausibly count as values under one or more of the concepts discussed in Section 2. A notable example is Ohkruhlik (1994), who focuses on "social factors."

Norms	Climate	Opinions	Interests	
Maxims	Professional culture	Worldviews	Virtues	
Desiderata	Decision vectors	Ideologies	Norms	
	Motivated reasoning (e.g., heuristics)	Background beliefs or assumptions	Maxims	
	Utilities	Interests	Desiderata	
	Virtues	Utilities		
	Worldviews	Preferences		
	Identities	Goals		
	Positionality	Aims		
	Ideologies	Purposes		
	Background beliefs or assumptions			

Recall that the first concept from Section 2 characterized values as criteria for making choices in science. We saw that Kuhn (1977) and McMullin (1983) characterized these criteria as "values" because they were not treated as hard-and-fast rules but rather as desiderata that needed to be weighed against each other. Thus, other phenomena that have this quality of guiding choice in a non-rule-governed fashion could plausibly be characterized as values under this concept. Concepts that are sometimes used to capture these sorts of phenomena include standards, guidelines, virtues, norms, maxims, and desiderata. As already noted, some of the cases in which it would not make as much sense to regard these concepts as falling under the umbrella of values would be when they are employed in a more rigid, rule-governed way. For example, Douglas (2013) argues that some cognitive values, such as internal consistency and empirical adequacy (relative to existing evidence), operate as "minimal criteria" for adequate science. Because these criteria must always be met, she questions whether they should be labeled as "values" (Douglas 2009, 94).

A particularly wide variety of concepts could fall under the second concept from Section 2 (i.e., factors that play a causal role in scientific choices, habits, or practices). Some of these concepts include biases, emotions, interests, organizational climate, professional culture, decision vectors, motivated reasoning [e.g., heuristics], opinions, virtues, worldviews, identities, positionality, ideologies, and background beliefs or assumptions (see Table 2). We will not attempt to define all these entities here because it is relatively clear that they are the sorts of things that can exert subtle or overt causal influences on scientists' reasoning or practices, but for those who want to read further about some of these phenomena we recommend sources like Douglas and Elliott (2022), Hilligardt (2022), Pennock and O'Rourke (2017), Schneider et al.

(2013), and Solomon (2001). As we discussed in Section 2, there is room for discussion about whether all psychological factors that causally influence scientists' choices should actually be considered values. For example, one might think that it is a stretch to label biases, heuristics, or other forms of motivated reasoning as "values" because they do not seem sufficiently closely linked to people's views about what is desirable. However, given the diverse contexts in which people discuss values in science, we are more inclined to discuss the strengths and weaknesses of placing various concepts under the umbrella of "values" rather than definitively concluding whether or not they should count as values.

Turning to beliefs or attitudes about what is desirable (the third concept of values from Section 2), some of the phenomena that could operate as causal factors could also sometimes fall under this concept, especially when they are consciously endorsed in some fashion. These include phenomena like worldviews, ideologies, and background beliefs or assumptions. In most cases, these phenomena would include a variety of elements, and beliefs about what is desirable might constitute one of those elements. Thus, it would probably be more accurate to say that these phenomena *incorporate* values. Other concepts that seem to fall under this concept of values include interests, utilities, preferences, aims, goals, and purposes. At first glance, utilities and preferences appear to fall squarely under this concept insofar as they constitute attitudes about what is desirable. What might make them somewhat more difficult to characterize is that they are not always accepted in a conscious fashion; sometimes they must be inferred from people's behaviors. Even when they are merely inferred from behaviors, however, one might still regard these as attitudes about what is desirable, especially if they would be endorsed if brought to light and scrutinized.[10] Other phenomena, such as aims, goals, and purposes, could arguably fall into either this third concept of values or the fourth one (desirable things in themselves), depending on whether one were to interpret them as beliefs about what is desirable or as desirable things themselves.

When aims, goals, and purposes are interpreted as desirable things in themselves rather than as beliefs about those things, they would fall under our fourth concept of values. One might also classify virtues under this concept, insofar as they consist of habitual forms of behavior that are desirable. Similarly, one might even classify norms or maxims as desirable things, although their desirability would presumably arise indirectly because of their usefulness for guiding people's behavior in socially beneficial ways. Whereas almost everyone would regard beliefs or attitudes about desirable things as values, it is somewhat less clear that everyone would regard desirable things in themselves as values. Thus, whether or not one would regard the phenomena that fall under this fourth concept as values would depend on one's overall views about this concept.

Background beliefs and assumptions are particularly complex concepts that merit a bit more discussion. We think that they can play a variety of different roles in science, and therefore they exhibit varying relationships to values. In some cases, background beliefs or assumptions could themselves serve as values in a manner that is relatively uncontroversial. For example,

---

[10] The authors would like to thank Zina Ward for helpful discussions about these issues, but the responsibility for any confusions or errors remains with the authors.

one might hold the background assumption that promoting public health is desirable, which would be a belief about what is desirable that falls into the third concept from Section 2. In other cases, background assumptions could serve as values in the sense of causal factors. For example, the background assumption that men are more socially important than women might cause scientists to put more effort into investigating issues that primarily affect men as opposed to issues that primarily affect women. In this case, some people might regard this background assumption as a value, whereas others might prefer to distinguish causal factors like this from genuine values.

In other cases, a background assumption might fall outside any of the concepts from Section 2, but it could still be related to values in important ways. For example, consider debates about how to interpret the evidence that exposure to very low doses of some chemicals could have endocrine-disrupting properties. Some scientists hold the background assumption that chemicals are unlikely to display harmful biological effects at low doses that differ from their effects at higher doses (Elliott 2011). Other scientists reject this background assumption and hold that it is perfectly reasonable that chemicals could have harmful effects at low dose levels while not having those effects at higher dose levels. One might not classify this background assumption itself as a value, but acceptance of it in the context of the contemporary chemical regulatory environment tends to promote achieving desirable states of affairs like public health (i.e., a value that falls under the fourth concept). Thus, one might say that this background assumption "encodes" or "supports" or "mediates" values. A similar "encoding" relationship could arise if concerns for public health played a causal role (i.e., acting as values under the second concept) in encouraging scientists to accept the background assumption.

The upshot of our analysis in this section is that values relate to other concepts in many different ways. Some concepts fall pretty straightforwardly under the umbrella of values. For example, insofar as utilities or preferences count as attitudes about what is desirable, they clearly count as values. Other concepts, such as worldviews or ideologies, can *incorporate* values, even if they are not identical to values, insofar as they include beliefs about what is desirable. Still other concepts, like identities or biases, typically fall under the concept of values as factors that causally influence scientists and their practices, and so they might or might not count as values, depending on whether one were inclined to count all factors that causally influence scientific reasoning and practice as genuine values. Finally, entities like background beliefs/assumptions could act as values in some cases (i.e., when they are beliefs about what is desirable) while having other relationships with values (e.g., "encoding" or "supporting" them) in other cases.

Given this analysis, it should be no wonder that the relationships between values and other concepts have been difficult to untangle. Because not everyone would regard all four concepts of values from Section 2 as genuine descriptions of values, there can be disagreement about whether particular concepts fall under the umbrella of values. Additional complications arise because many of the concepts that we have considered (e.g., organizational climate, professional culture, worldviews, identities, aims, goals, purposes) capture phenomena that can potentially fall under multiple concepts from Section 2, depending on the context and the ways

the concepts are interpreted. For example, aims, goals, and purposes could be regarded as beliefs or attitudes about what is desirable (our third concept), or they could be regarded as desirable things themselves (our fourth concept). If one regarded desirable things themselves as genuine values, then one would presumably conclude that aims, goals, and purposes consistently fall under the umbrella of values. However, if one regarded only beliefs or attitudes about desirable things as values and not desirable things themselves, then one would have to conclude that goals, aims, and purposes count as values under some usages but not others. Thus, although the brief analysis in this section hardly settles all the questions about how values relate to other concepts, it does suggest a path forward for thinking more carefully about them.

4. Clarifying the Science and Values Literature

Although we think that the clarification of values and related concepts in the preceding sections of this paper is important in its own right, an added benefit is that it has the potential to bring greater clarity to the science and values literature more broadly. One of the most important insights stemming from the analysis in Section 2 is that, once one distinguishes our four concepts of values, it becomes clear that different concepts of values tend to be at play in different contributions to the science-and-values literature. However, these different concepts have not been clearly recognized because the same term (i.e., 'values') has been used to refer to different things. This section explores how different concepts of values are present in two areas of the literature: arguments against the VFI and proposals for managing values. We then highlight the practical upshot of this analysis–namely, that the literature on values and science should be viewed less as a unified enterprise focused on a single phenomenon and more as a patchwork of arguments and approaches that address related but nevertheless distinct phenomena. As noted previously, we think it is reasonable (and even desirable) to continue exploring an array of different concepts and phenomena under the umbrella of "values and science," but it is important to clarify which ones are under investigation.

Arguments against the VFI

In order to illustrate how different concepts of values are at play in different arguments against the VFI, let us consider the four arguments that one of us described in a recent overview of the literature on science and values (Elliott 2022). The overview begins with the "gap" argument: "Evidential gaps between data and conclusions are inevitably filled by value-laden background assumptions, and thus, it does not make sense to exclude non-epistemic values from scientific reasoning" (Elliott 2022, 20; see e.g., Howard 2009; Longino 1990; Biddle 2013). Rather than pursuing the quixotic goal of trying to exclude non-epistemic values from scientific reasoning, proponents of this argument typically argue that it is best to evaluate background assumptions (and the values associated with them) critically from a diversity of perspectives. This argument is deeply influenced by the work of Helen Longino (1990), who argued that background assumptions are necessary in order to establish evidential relationships between data and hypotheses. According to Longino, these background assumptions are typically value-laden, either in the sense that values causally contribute to the acceptance of particular background assumptions or because the acceptance of those background assumptions

promotes the realization of particular values. Thus, this argument focuses most directly on a concept of values as causal factors (which contribute to the acceptance of background assumptions) or as desirable things (which can be affected by the acceptance of background assumptions).

The second argument against the VFI is the "error" argument: "When scientists face epistemic risks, they ought to factor the non-epistemic values at stake into their decisions about how to navigate the risks" (Elliott 2022, 23). Douglas's classic version of the error argument focuses on inductive risk, where scientists face the risk of drawing a false negative or a false positive conclusion (see e.g., Douglas 2009; Douglas 2021; Elliott and Richards 2017), but it could also involve representational risks (Biddle and Kukla 2017; Harvard and Winsberg 2022). Douglas argues that scientists have ethical responsibilities to consider the potential consequences of drawing erroneous conclusions so that they do not negligently or recklessly cause harm to others as a result. But this argument justifies including a much narrower range of "values" in scientific reasoning than the gap argument does. It focuses on ethical values about what risks are acceptable or unacceptable to accept. These values appear to fall primarily in the third category from Section 2, i.e., beliefs or attitudes about what is desirable. Significantly, the error argument clearly does not justify incorporating other causal factors in scientific reasoning besides ethical considerations about how to weigh the consequences of error.

The third, "aims," argument focuses on values in other senses. The argument states that "in order to achieve the non-epistemic aims of science, scientists need to take nonepistemic values into account when assessing the quality of scientific models, hypotheses, and theories" (Elliott 2022, 29). A significant benefit of the distinctions provided in Section 2 is that they facilitate a more careful analysis of the concepts associated with this argument. First, the "aims" of science arguably count as values in either the third or fourth senses discussed in Section 2. In other words, they constitute desirable states of affairs or beliefs about the desirability of those states of affairs. Proponents of the aims argument argue that hypotheses, theories, and models should be evaluated based on the extent to which they display characteristics that promote the realization of those desirable states of affairs, even when those characteristics are not purely epistemic in character. For example, in order to promote the aim of regulating hazardous chemicals effectively, Elliott and McKaughan (2014) argue that it is appropriate to evaluate risk assessment methods based on whether they have characteristics (e.g., speed and ease of use) that will promote that aim (see also Lusk and Elliott 2022). Those characteristics count as values in the first sense discussed in Section 2; namely, they are properties of hypotheses, theories, or models that count as criteria for evaluating them. Thus, the aims argument justifies incorporating particular sorts of values (i.e., criteria) by appealing to the way they help to achieve other sorts of values (i.e., desirable states of affairs and/or beliefs about them). Like the error argument, this approach justifies including only a relatively circumscribed array of phenomena in science under the label of "values."

The fourth, "conceptual," argument states that "non-epistemic values are relevant to assessing scientific hypotheses that incorporate 'mixed' or value-laden concepts" (Elliott 2022,

31).[11] To flesh out the notion of mixed concepts, Elliott appeals to the work of Anna Alexandrova (2018), who discusses concepts like well-being and insists that they cannot be adequately characterized apart from ethical value judgments about what account of well-being is most justifiable. A slightly different version of the conceptual argument appeals to the social consequences of terminological choices in science and the ethical importance of reflecting on those consequences (see Elliott 2022). Either way one interprets the argument, however, it appeals to ethical judgments, which fall into the third concept of values from Section 2 (i.e., beliefs or attitudes about what is desirable).[12]

This clarification of the kinds of values associated with different arguments against the VFI shows that they do not actually justify incorporating the same entities in scientific reasoning. The error and conceptual arguments mainly support forms of ethical reasoning about how to act as scientists in a socially responsible manner. In the case of the error argument, the ethical reasoning at issue concerns the best ways of handling epistemic risks (such as the risk of drawing false positive or false negative conclusions or the risk of choosing inappropriate representations; see Harvard and Winsberg 2022). In the case of the conceptual argument, it involves assessing the most appropriate ways to interpret "thick" concepts and the most appropriate conceptual schemes to employ. In contrast, the aims argument focuses not so much on ethical reasoning but rather on employing criteria that help to advance the aims of inquiry, even when those criteria are non-epistemic in character. The consideration of these criteria when assessing scientific theories or models differs from the sorts of ethical reasoning considered under the error and conceptual arguments.[13] Finally, the gap argument is particularly different in character from the other three. It is far less exclusive about the kinds of "values" that it justifies incorporating in scientific reasoning; it allows for causal factors in general to influence scientific reasoning as long as those factors are subjected to appropriate critical scrutiny.

Management Strategies

---

[11] In some cases, the distinction between the conceptual argument and the error argument could break down. For example, in some cases of representational risk, scientists could face the risk of adopting a conceptual framework that misrepresents the features of a phenomenon that it ought to represent. This would plausibly qualify as a case where both the error argument and the conceptual argument are applicable. These arguments do not always coincide, however. For example, there could be cases where a conceptual choice does not count as an erroneous representation, but it is nevertheless ethically significant and thus subject to the conceptual argument.

[12] One could potentially also interpret this argument as involving the first concept of values (i.e., criteria for choice). For example, one might adopt particular criteria or standards for assessing well-being. Our inclination, though, is that those criteria or standards should ultimately be based on ethical beliefs (i.e., the third concept of values).

[13] Admittedly, the aims argument still faces the question of how to justify the aims of inquiry. Depending on one's preferred approach, the justification process might include ethical reasoning about what aims are most appropriate. Nevertheless, it is crucial to recognize that this ethical reasoning would involve choosing the *aims* of inquiry and not the *criteria* for assessing scientific theories or models. For example, Elliott and Lusk (2022, 17) discuss a scenario in which scientists need to evaluate chemical risk assessment methods based on their ability to promote effective regulation of toxic chemicals. The criteria for evaluating these methods are likely to be grounded in means-ends reasoning about whether the criteria are likely to promote a particular conception of effective regulation that is already in place, not on deeper reasoning about what is ultimately desirable.

Once one begins to distinguish the different concepts of values from Section 2, it becomes clear that different strategies for managing values also tend to be directed at some concepts rather than others. Consider, for example, the five general kinds of management strategies discussed by Holman and Wilholt (2022). Their first, "axiological," strategy calls for identifying a set of values that are appropriate for influencing science, perhaps because they can be justified through ethical analysis or political deliberation. This strategy makes a great deal of sense if one is thinking about values in the third sense discussed in Section 2 (i.e., beliefs or attitudes about what is desirable). For example, one might engage in ethical or political analysis to determine which values should guide the assessment of epistemic errors or the choice of scientific concepts. Ethical or political analysis could also help in assessing which aims should guide scientific inquiry. However, it makes much less sense to apply this management approach if one is thinking about values in the sense of causal factors because those factors are not always subject to conscious consideration and analysis.

Holman and Wilholt's second, "functionalist," approach to managing values focuses on limiting them to specific uses or roles. For example, Douglas (2009) famously argued that values should play only indirect roles in scientific reasoning (i.e., assessing the amount of evidence needed to draw conclusions) rather than direct roles (i.e., counting as reasons or evidence). Douglas designed her functionalist approach to apply specifically in the context of the error argument, where it serves as a strategy for managing values that operate as beliefs about what is desirable.[14] However, as in the case of the axiological management approach, the functionalist approach does not appear to be as helpful for handling values when they are interpreted as causal factors because they cannot always be limited and subjected to conscious analysis. In addition, when values operate as criteria for assessing scientific theories or methods in order to achieve non-epistemic aims, it might not be necessary or appropriate to limit them to only a specific functional role, at least in the sense that Douglas envisioned (see e.g., Elliott and McKaughan 2014).

Holman and Wilholt's (2022) third and fourth approaches for managing values are "consequentialist" and "coordinative." The consequentialist management strategy assesses whether value influences in science are appropriate based on whether they help to generate particular consequences, and the coordinative strategy assesses value influences based on whether they align with the expectations or goals of the users or recipients of scientific information. These approaches work particularly well when values are conceptualized as criteria for choice because it makes sense to assess those criteria based on the consequences of employing them or their alignment with the aims of inquiry. This approach largely takes particular aims as given, however, and thus it is less helpful as a strategy for managing values when they are interpreted as those aims themselves (i.e., the third or fourth categories from

---

[14] It is important to recognize that Douglas does not rely solely on a functional approach to managing values (see e.g., Douglas 2021), but our point is merely that a functional approach could potentially make sense as one of the tools for managing values in the context of the error argument, whereas it would not make sense in other contexts.

Section 2). Instead, one would presumably need to engage in some sort of ethical or political reasoning (as in the axiological management strategy) to assess those aims.[15]

Finally, Holman and Wilholt's (2022) "systemic" management approach strives to create a community structure that assesses value influences and weeds out problematic ones. This is arguably the management approach that makes the most sense when dealing with causal factors (especially those that do not seem particularly amenable to ethical or political justification, such as features of a scientist's identity or positionality). Thus, given that Longino tended to think about values in the form of causal factors, it makes sense that she developed a systemic approach for managing them (see e.g., Longino 1990; Longino 2002). Of course, this is not to say that a systemic approach is applicable *only* to values interpreted as causal factors; it could obviously be helpful to develop a system of community assessment for managing values interpreted in other ways as well. However, when addressing other concepts of values, such as beliefs or attitudes about what is desirable, the systemic approach might need to be supplemented with additional strategies for "weeding out" values that are too problematic to be included in community deliberations (see e.g., Intemann 2017; Schroeder 2021).

The Upshot

The significance of clarifying the different concepts of values at play in different arguments against the VFI and in different management strategies is that it reveals ways in which the literature on values and science is less unified than it might otherwise appear to be. Because the common language of "values" is used throughout the literature, it gives the impression that the same basic ideas are at play throughout. Once one clarifies the different concepts being used, however, it turns out that the literature looks much more like a patchwork of discussions about related but distinct phenomena. To illustrate this point, consider three distinct activities that are discussed as part of this patchwork: (1) recognizing and managing factors that causally influence scientists' reasoning; (2) ethically assessing epistemic risks and conceptual choices; and (3) pursuing non-epistemic aims in science. Although they are typically discussed together under the rubric of "values in science," attending to the distinctions between different concepts of values highlights significant differences between these three areas of discourse.

The first area of discourse is deeply influenced by the work of Longino (1990; 2002). As seen above, she worked largely with a concept of values as factors that causally influence scientists' reasoning and practices, and she was willing to consider a wide array of causal factors under this label because her primary interest was in the ways that background assumptions could subtly encode or transmit their influences. As noted in Section 3, this encoding relationship could occur either through these factors causing scientists to accept particular background assumptions or through the acceptance of those background assumptions having effects on other desirable things. She thought that the best approach for

---

[15] One could, of course, assess the aims of a specific inquiry based on the extent to which they help to achieve others consequences or aims. However, this would seemingly result in a regress unless one ultimately arrived at some other way to assess those deeper consequences or aims.

managing all these causal factors (which often went unrecognized) was to bring together a diverse array of interlocutors who could help to "unearth" them and critically evaluate them from an array of different perspectives.

The second area of discourse involves the ethical assessment of epistemic risks and conceptual choices. To understand the nature of this discourse, one can focus on Douglas's work on inductive risk (e.g., Douglas 2009; Douglas 2021; but see also Biddle and Kukla 2017; Harvard and Winsberg 2022). Although the work of both Longino and Douglas are typically treated together under the label of "values," when one stops to consider the entities at play in their work, they are very different.[16] Whereas Longino focused on a wide array of causal factors, Douglas focused strictly on ethical judgments about the sorts of harmful outcomes that scientists have responsibilities to avoid. It is no wonder, then, that she advocates a different management approach for handing values because she is addressing very different sorts of phenomena. It would not make sense to employ a functionalist approach in response to Longino's causal factors because they can affect scientific reasoning in a wide variety of ways, but it is more feasible to strive for a functionalist approach for managing Douglas's ethical values because they have a more narrow role to play in scientific reasoning (i.e., assessing standards of evidence).[17] Douglas's approach to values also generates a different sense in which scientific choices can be regarded as "value laden" (see Ward 2021): in contrast to Longino's focus on values causally affecting or being affected by choices in science, Douglas's ethical values serve as motivating and justifying reasons for choices.

Finally, the third area of discourse draws from figures like Elliott (2013), Intemann (2015), and Hicks (2022), who think in terms of the aims of science. Although they also talk about "values," they are primarily thinking about the goals or purposes of inquiry and the criteria required for assessing scientific work in order to achieve those goals or purposes. These "values" (both the criteria and the goals they help to achieve) are distinct both from Longino's causal factors and from Douglas's ethical judgments. Not surprisingly, the preferred approach for managing these values is also different: a coordinative approach makes sense in this context because the criteria for assessing scientific work need to align with the goals or purposes of inquiry. And the sense in which scientific choices turn out to be value-laden in this discourse coincides with the nature of these values as well. Like Douglas, those who focus on the aims of science are interested in the ways that values (i.e., criteria) can serve as motivating or justifying reasons for particular choices. In addition, they are interested in how particular criteria end up being value-laden in the sense that they generate socially relevant effects (see Ward 2021).

To reason appropriately about values, even in a single scientific context, these different discourses about values often need to be distinguished. For example, in the context of chemical risk assessments, scientists are often subject to both financial and non-financial conflicts of

---

[16] But see ChoGlueck (2018) for more discussion of the relationships between Longino's and Douglas's work.

[17] Even though a functionalist approach is more appropriate in the context of Douglas's work than in Longino's, the feasibility of Douglas's particular functionalist approach has still been challenged (e.g., Bluhm 2017; Elliott 2013; Steel and Whyte 2012).

interest (COIs), which could operate as causal factors that influence their reasoning in subconscious ways. These values, which fit the first discourse above, might often be unhelpful or even harmful to science if they are not appropriately "balanced" (see e.g., Solomon 2001), and it might even make sense to exclude individuals with severe COIs from participating in some risk-assessment activities (Resnik 2023). At the same time, in order to do their work in a socially responsible fashion, the scientists involved in chemical risk assessment need to reflect on how much evidence to demand before drawing conclusions for policy makers and members of the public. This activity involves values operating in accordance with the second discourse discussed above. To manage these values requires ethical or political procedures for selecting appropriate values (Schroeder 2022) and perhaps limiting their influences to particular aspects of the risk-assessment process (Douglas 2009). Finally, when doing their work for regulatory purposes, risk assessors might also need to evaluate their methodological choices by considering which ones are most appropriate for meeting specific regulatory goals or requirements. This involves values operating as criteria for achieving aims, in accordance with the third discourse about values. To manage these values requires means-ends reasoning about the extent to which specific criteria promote particular aims, and (when feasible) reflecting on which regulatory goals are ultimately appropriate.

What drives home the patchwork nature of these different discourses is the fact that one could adopt a positive attitude toward the kinds of values associated with some of them while resisting the kinds of values associated with others. For example, one could affirm the importance of assessing epistemic risks and conceptual choices (and deliberately bring values into scientific reasoning in that sense) while at the same time trying to minimize the influence of most other causal factors in science (thereby resisting the inclusion of values in that sense). Or one could resist the notion that scientists should bring values into their work when assessing epistemic risk (see e.g., Betz 2013; Betz 2017) while accepting that values should play a role as non-epistemic criteria that can help to achieve the aims of science, even when they are at least partly non-epistemic in character. Alternatively, one could accept that scientists should bring ethical values into their assessments of epistemic risk while questioning the extent to which non-epistemic criteria should be employed as a basis for assessing scientific hypotheses or models. In other words, one could reject the "value-free ideal" for some kinds of values while accepting the "value-free ideal" for other kinds of values (Elliott forthcoming). Although this point might seem obvious in retrospect, prominent discussions of the literature on values and science have often merged these discourses together when formulating criticisms of a single value-free ideal (see e.g., Elliott 2022). Moving forward, it seems clear that discussions of values in science should clarify the kinds of values at play in order to avoid generating confusion about what kind of discourse is at issue and which kinds of values are actually being defended and managed.

Conclusion

The literature on values and science has flourished in recent years. It is rich, dynamic, and—as we have argued here—more diverse than is often recognized. We have distinguished four different concepts of values: (1) values as criteria or standards; (2) values as psychological

factors that causally influence scientists' reasoning; (3) values as beliefs or attitudes about what is desirable; and (4) values as desirable things themselves. There are complex relationships between these different concepts, and they each have strengths and weaknesses. We suggest that our analysis can help researchers not just to be more precise but also to be more thoughtful about choosing concepts that accord with their goals and with the contexts in which they work.

We have also shown how our four-fold framework can help to clarify the nature of other, related concepts because one can see how they fall into one or more of these concepts or how they bear some other relationship to values. For example, we argued that background assumptions could act as values in some cases (i.e., when they are beliefs about what is desirable) while having other relationships with values (e.g., "encoding" or "supporting" them) in other cases.

Finally, we argued that this conceptual clarification can advance the science-and-values literature by showing that many contributions to the literature (including arguments against the value-free ideal and strategies for managing values) have worked with different concepts of values while failing to make these differences clear. Attention to these differences suggests that the literature on values in science should be viewed as a patchwork of discourses that are related but that should not be conflated simply because of the common terminology of "values." In our view, this literature can continue to be a broad tent that encompasses a variety of different concepts, but we think this literature will move forward more productively if there is greater clarity about the particular concepts of values at play.

The clarifications that we have provided here suggest a number of avenues for further scholarship. There is still much to be said about how our four concepts of values should be defined and how they relate to one another and to other concepts. And, as we discussed briefly in Section 4, attending to the different concepts of values that we have examined in this paper is also likely to shed further light on arguments for and against the cogency of the VFI. For example, the VFI is likely to appear much more compelling if one is interpreting values as extraneous factors that causally influence scientists in idiosyncratic ways rather than as ethical reasons for handling underdetermined choices in socially responsible ways (see e.g., Douglas and Elliott 2022; Elliott forthcoming; Hudson 2021). More broadly, our categorization of four different concepts of values might help elucidate the often-contested relationships between epistemic and non-epistemic values. Epistemic values have traditionally been discussed primarily in the context of our first concept of values (i.e., criteria for choice), while non-epistemic values can fit into all four of our concepts, depending on the context. Perhaps attention to these differences could help to clarify different functional roles that values can play. For example, one could distinguish values that play the functional role of serving as criteria for theory choice from values that serve the functional role of serving as outcomes that would be desirable to achieve. These functional distinctions could serve as a useful successor to the epistemic/non-epistemic distinction even if the criteria for theory choice cannot be cleanly distinguished into those that are epistemic and those that are not. Moving forward, we are hopeful that these sorts of efforts to clarify the fundamental concept of "value" will contribute substantially to the vibrant and socially important literature on values and science.

References

Alexandrova, A. 2018. Can the science of well-being be objective? *British Journal of Philosophy of Science* 69: 421-445.

Anderson, E. 1993. *Values in Ethics and Economics.* Cambridge, MA: Harvard University Press.

Betz, G. 2013. In defence of the value free ideal. *European Journal for Philosophy of Science* 3: 207-220.

Betz, G. 2017. Why the argument from inductive risk doesn't justify incorporating non-epistemic values in scientific reasoning. In K. Elliott and D. Steel (eds.), *Current Controversies in Values and Science*, pp. 94-110. London: Routledge.

Biddle, J., 2013. State of the field: Transient underdetermination and values in science. *Studies in History and Philosophy of Science* 44: 124-133.

Biddle, J. and Kukla, R. 2017. The geography of epistemic risk. In K. Elliott and T. Richards (eds.), *Exploring Inductive Risk: Case Studies of Values in Science,* pp. 215-238. New York: Oxford University Press.

Bluhm, R. 2017, Inductive risk and the role of values in clinical trials. In K. Elliott and T. Richards (eds.), *Exploring Inductive Risk: Case Studies of Values in Science,* pp. 193-212. New York: Oxford University Press.

Brown, M. 2020. *Science and Moral Imagination: A New Ideal for Values and Science.* Pittsburgh: University of Pittsburgh Press.

ChoGlueck, C. 2018. The Error Is in the Gap: Synthesizing Accounts for Societal Values in Science. *Philosophy of Science* 85: 704–25.

Dietz, T. 2013. Bringing values and considerations to science communication. *Proceedings of the National Academy of Sciences* 110: 14081-14087.

Douglas, H. 2009. *Values, Policy, and the Value-Free Ideal.* Pittsburgh: University of Pittsburgh Press.

Douglas, H. 2021. *The Rightful Place of Science: Science, Values, and Democracy. The 2016 Descartes Lecture*s, ed. Tempe, T. Richards: Consortium for Science, Policy, and Outcomes.

Douglas, H. and Elliott, K. 2022. Addressing the reproducibility crisis: A response to Hudson. *Journal for General Philosophy of Science* 53: 201-209.

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

Elliott, K. 2013. Douglas on values: From indirect roles to multiple goals. *Studies in History and Philosophy of Science* 44: 375-383.

Elliott, K. 2022. *Values in Science.* Cambridge: Cambridge University Press.

Elliott, K. Forthcoming. "Characterizing the Value-Free Ideal: From a Dichotomy to a Multiplicity," in J. Tsou, J. Shaw, and C. Fehr (eds.), *Values, Pluralism, and Pragmatism: Themes from the Work of Matthew J. Brown.* Boston Studies in the Philosophy and History of Science. Cham: Springer.

Elliott, K and McKaughan, D. 2014. Nonepistemic values and the multiple goals of science. *Philosophy of Science* 81: 1-21.

Elliott, K. and Richards, T. 2017. *Exploring Inductive Risk: Case Studies of Values in Science.* New York: Oxford Press.

Harvard, S. and Winsberg, E., 2022. The epistemic risk in representation. *Kennedy Institute of Ethics Journal 32*: 1-31.

Hicks, D. J. 2022. When Virtues are Vices: 'Anti-Science' Epistemic Values in Environmental Politics. *Philosophy, Theory, and Practice in Biology* 14*: 1-30.*

Hilligardt, H. 2022. Looking beyond values: The legitimacy of social perspectives, opinions, and interests in science. *European Journal for Philosophy of Science* 12: 1-20.

Holman, B. and Wilholt. T. 2022. The new demarcation problem. *Studies in History and Philosophy of Science* 91: 211-220.

Howard, D. 2009. Better red than dead–Putting an end to the social irrelevancy of postwar philosophy of science. *Science and Education* 18: 199-220.

Hudson, R. 2021. Should we strive to make science bias-free? A philosophical assessment of the reproducibility crisis. *Journal for General Philosophy of Science* 52: 389-405.

Intemann, K. 2015. Distinguishing between legitimate and illegitimate values in climate modeling. *European Journal for Philosophy of Science* 5: 217-232.

Intemann, K. 2017. Feminism, values, and the bias paradox: Why value management is not sufficient. In K. Elliott and D. Steel (eds.), *Current Controversies in Values and Science*, pp. 130-144. London: Routledge.

Karwat, M. 1982. Political values as ideas of social needs. *International Political Science Review* 3: 198-204.

Kuhn, T. 1977. Objectivity, Value Judgement, and Theory Choice. In Kuhn, T. (ed.), *The Essential Tension*, 320–9. Chicago: University of Chicago Press.

Lacey, H. 1999. *Is Science Value Free? Values and Scientific Understanding*. New York: Routledge.

Lacey, H. 2017. Distinguishing between cognitive and social values. In K. Elliott and D. Steel (eds.), *Current Controversies in Values and Science*, p. 15-30. London: Routledge.

Longino, H. 1990. *Science as Social Knowledge.* Princeton: Princeton University Press.

Longino, H. 2002. *The Fate of Knowledge*. Princeton: Princeton University Press.

Lusk, G. and Elliott, K. 2022. Non-epistemic values and scientific assessment: An adequacy for purpose view. *European Journal for Philosophy of Science* 12*:* 1-22.

McMullin, E. 1983. 'Values in Science'. In Asquith, P. and Nickles, T. (eds.), *PSA 1982*, 3–28. East Lansing: Philosophy of Science Association.

Pennock, R.T. and O'Rourke, M., 2017. Developing a scientific virtue-based approach to science ethics training. *Science and Engineering Ethics* 23: 243-262.

Resnik, D. 2023. Disclosing and managing non-financial conflicts of interest in scientific publications. *Research Ethics* 19: 121-138.

Schneider, B., Ehrhart, M.G. and Macey, W.H., 2013. Organizational climate and culture. *Annual Review of Psychology*, *64*, pp.361-388.

Schroeder, S. A. 2021. Democratic values: A better foundation for public trust in science. *British Journal for the Philosophy of Science* 72: 545-562.

Schroeder, S. A. 2022. Thinking about vaues in science: Ethical versus political approaches. *Canadian Journal of Philosophy* 52: 246-255.

Solomon, M. 2001. *Social Empiricism.* Cambridge, MA: The MIT Press.

Solomon, M., 2012. Socially responsible science and the unity of values. *Perspectives on Science* 20: 331-338.

Steel, D. 2010. Epistemic values and the argument from inductive risk. *Philosophy of Science, 77*(1), pp. 14-34.

Steel, D. and Whyte, K. 2012. Environmental justice, values, and scientific expertise. *Kennedy Institute of Ethics Journal 22*(2), 163-182.

Swift, A. 2001. *Political Philosophy: A Beginners' Guide for Students and Politicians.* Cambridge: Polity Press.

Vaske, J. and Manfredo, M. 2012. Social psychological considerations in wildlife management. In D. Decker, S. Riley, and W. Siemer (eds.), *Human Dimensions of Wildlife Management,* pp. 43-57. Baltimore, MD: Johns Hopkins Press.

Ward, Z. 2021. On value-laden science. *Studies in History and Philosophy of Science* 85: 54-62.

Winsberg, E. 2024. Managing values in science: A return to decision theory. Unpublished manuscript.

Zhao, H. 2022. The nature of values in science: What they are and how they guide. Unpublished manuscript.

Zhao, H. 2023. Science in a Sexist Society. Ph.D. diss, Columbia University.