ORIGINAL RESEARCH



Values as heuristics: a contextual empiricist account of assessing values scientifically

Christopher ChoGlueck¹ · Elisabeth A. Lloyd²

Received: 20 July 2022 / Accepted: 19 May 2023 / Published online: 13 June 2023 © The Author(s) 2023

Abstract

Feminist philosophers have discussed the prospects for assessing values empirically, particularly given the ongoing threat of sexism and other oppressive values influencing science and society. Some advocates of such tests now champion a "values as evidence" approach, and they criticize Helen Longino's contextual empiricism for not holding values to the same level of empirical scrutiny as other claims. In this paper, we defend contextual empiricism by arguing that many of these criticisms are based on mischaracterizations of Longino's position, overstatements of certain claims, and false dichotomies. Her contextual empiricism not only allows for the empirical support and disconfirmation of values, but Longino explicitly discusses when values can be empirically adjudicated and emphasizes the crucial role of the community for standards of evidence. We support contextual empiricism and elaborate a less direct account of "values as heuristics" by reviewing Longino's theory of evidence and then using a case study from Elisabeth Lloyd on the biology of female orgasm, demonstrating the disconfirmation of androcentric values in evolutionary science. Within Longino's and Lloyd's contextual empiricism, values do not get treated as empirical evidence to be directly assessed by individuals, but rather values are heuristic tools to build models whose use can be validated or invalidated by communities based on their empirical fruitfulness in the logic and pragmatics of research questions in specific historical and cultural contexts.

Keywords Values as evidence \cdot Contextual empiricism \cdot Androcentrism in biology \cdot Underdetermination \cdot Community standards of evidence \cdot Feminist philosophy of science

Christopher ChoGlueck christopher.choglueck@nmt.edu

¹ Department of Communication, Liberal Arts, and Social Sciences, New Mexico Tech, Socorro, NM, USA

² Department of History and Philosophy of Science and Medicine, Indiana University, Bloomington, IN, USA

1 Introduction

Traditionally, empiricists have held values in suspicion (or outright contempt), following Hume's classic argument that normative conclusions cannot be derived from purely empirical observations (Marchetti & Marchetti, 2016). Yet, with more attention to the contextual nature of science, philosophers have shown that values have guided theory choice and, conversely, that widely held scientific values like accuracy, consistency, scope, and fruitfulness are informed by scientists' experiences of success and failure (Kuhn, 1977).

Moreover, because of their normative commitments, empiricists of a more feminist stripe like Helen Longino (1990) and Elisabeth Lloyd (1995) have led the vanguard in collapsing this fact-value dichotomy by challenging idealized concepts of "value-free" or "value-neutral" science and articulating value-rich conceptions of scientific objectivity (Crasnow, 2013; Intemann, 2010; Richardson, 2010). With their alternative value frameworks, feminist scientists like biologist Ruth Hubbard (1979), neuro-physiologist Ruth Bleier (1984), and sociologist Patricia Hill Collins (1986) have successfully challenged dominant oppressive values in science like sexism, racism, and heteronormativity (Schiebinger, 2001). Following these insights into how values guide and infuse empirical inquiry—for bad *or* for good—feminist philosophers have more recently turned to the reverse direction: how is it that scientific facts inform our values, especially ethical and political values (Anderson, 2004; Clough, 1998; Yap, 2016)?

In reconsidering the empirical status of values, some feminists such as Sharyn Clough and Maya Goldenberg now advocate for treating "values as evidence" by holding value judgments "to be subject to the same level of rigorous empirical inquiry" as descriptive claims (Goldenberg, 2015, p. 25; Clough, 2013a). Goldenberg describes the current status of the field as such: "whereas scholarly attention has been focused on facts as value-laden and the impossibility of value-free science due to the normativity of experience, far less attention has been given to the facticity of values" (2021, p. 57, emphasis added). Surprisingly, the primary target of Clough and Goldenberg is feminist philosopher of science Helen Longino, on account of her analysis of underdetermination and objectivity and her claim that values "may not be subject to empirical confirmation or disconfirmation" (Longino, 1990, p. 75, quoted by Goldenberg, 2015, p. 12; see also Clough, 1998; Solomon, 2012). Is it true that Longino's contextual empiricism makes an exception for values (compared with descriptive claims)? Moreover, are values like feminism and androcentrism simply "objectively true" or "false" beliefs, meaning "value judgments, like any other [decision], just are empirical hypotheses, broadly speaking-hypotheses that can be subjected to rational processes of adjudication" (Clough & Loges, 2008, p. 88; contrast with Yap, 2016)? What can Longino's approach even offer us in support of feminism and social justice more generally?

There are huge stakes to these debates over the empirical status of values, given the power of science in society, the widespread influence of sexism and other oppressive values on science, and the adverse impacts of such science on society. On one end of the continuum is empirical optimism: if science can decisively adjudicate values as some more radical holists contend (Clough, 1998; Goldenberg, 2015), then scientists ought

to inform ethical decision-making and social policy-making in an even more direct and robust manner. While this could have significant social benefits, the deployment of such broad-strokes empiricism based on technical expertise could border on scientism and technocracy, both of which have tensions with democratic processes (Jasanoff, 1990). On the other end of the continuum is empirical pessimism: if science cannot assess values at all (the equally radical alternative view attributed to Longino), then we are back to the traditional fact-value dichotomy, in which values are either, at best, remote from the empirical world or, at worst, empirically immune to any scientific testing. Then, one might reasonably worry that facts about, e.g., women's lives and capacities, are compatible with both feminist and sexist values, rendering the facts insufficient for value revision (Alcoff, 2006).

In this paper, we articulate a middle-ground position between empirical optimism and pessimism, defending Longino's contextual empiricism on the abilities of science to assess values.¹ In contrast with the approach of "values as evidence," we emphasize the limitations of empirical "tests" of values (see also Solomon, 2012; Yap, 2016). We also elaborate how to assess values indirectly in terms of their empirical fruitfulness, where values operate as *heuristics* that provide frameworks for posing empirical questions, constructing models, and motivating the collection of certain kinds of evidence (see Longino, 2008). To do so, our analysis operates across levels of abstractness and concreteness. At the more conceptual level, we describe Longino's original program of contextual empiricism, particularly her views on background assumptions and evidential status. We contend that contextual empiricism not only allows for the empirical support/refutation of values but also that Longino explicitly discusses when values can be empirically adjudicated and when not.

Then, to link this theory with practice, we use a case study on gender bias in research on female orgasm from Elisabeth Lloyd (the second author on this present paper). Whereas Longino conceptualized how science must be done to combat hegemonic values, Lloyd's successful intervention in biology demonstrates how such processes actually work. Through this concrete case, we explore the prospects for a more tempered account of the normative power of science regarding the empirical status of values. We demonstrate how the empirical status of sexist values can be challenged by exposing their empirical deficiencies—and how the more empirically fruitful values from feminists provoked further empirical research to evaluate value-laden background assumptions, explore alternative theories, and provide the evidence needed for sexual liberation.

We begin in Sect. 2 with a close reading of Longino on the category of empirical evidence. Section 3 then evaluates criticisms of her treatment of values, showing how her critics at times mischaracterize her position, overstate certain claims, and rely on the rational-social dichotomy that Longino aimed to collapse (Longino, 2002). After elaborating the heuristic power of values, Sect. 4 details how oppressive values like androcentrism and heteronormativity can be empirically disconfirmed through what Longino calls "transformative criticism." This case study on the impacts of feminist

¹ While there are interesting questions about what sorts of values and what roles for those values are legitimate in science, these debates are beyond the scope of this paper (see Anderson, 2004; Douglas, 2009; Intemann, 2005, 2015; Holman & Wilholt, 2022).

heuristics throughout biology supports and extends Longino's framework on the complex, limited, but genuine invalidation of the use of certain values in specific contexts, emphasizing the importance of evaluating shared community standards empirically.

2 Longino on empirical evidence

Over three decades ago, Longino famously criticized the traditional "value-free" ideal of a dispassionate scientist detached from their personal values and biases, and she offered an alternative view of objectivity in science without value-freedom. Her methodological, interactive conception of objectivity is consistent with some influence of contextual values on science, based on the insight that effective and significant scientific criticism and reasoning about evidence can depend on value judgements, e.g., values implicit in dominant assumptions in that field (for other senses of objectivity, see Daston & Galison, 2007; Douglas, 2009; Porter, 1996; Crasnow, 2013; Lloyd, 1995; Lloyd & Schweizer, 2014). Values function properly in science as the grounds for "transformative criticism," so methodological objectivity depends on the diversity of values represented and the social structuring of science: recognized venues for criticism, shared public standards, community response (uptake of criticism), and tempered equality of intellectual authority (1990, p. 78f; 2002, pp. 128–135).

Now, what makes Longino's contextual-empiricist approach *empiricist* in the first place? Here, we explain her (1979) account of empirical evidence, upon which her later (1990) claims about objectivity and social relations stand.² The thing we call "evidence," according to Longino, is actually a *relation* between a fact (a state of affairs or "data") and a hypothesis, and that relationship is *extrinsic* to both the fact and the hypothesis:

What determines whether or not someone will take some fact or alleged fact, x, as evidence for some hypothesis, h, is not any natural (e.g., causal) relation between the state of affairs, x, and that described by h, but are other beliefs that person has concerning the evidential connection between x and h. (Longino, 1979, p. 37)

Those "other beliefs" connect the two: the relevance of a fact *as evidence* for a hypothesis necessarily involves additional "background beliefs or assumptions" (1979, p. 40). For instance: Why might you take red spots on your daughter's stomach to be evidence of measles? The *relevance* of this fact for that hypothesis might be based on the additional background belief that red spots are a *symptom* of measles. Such a connection, however, is both theoretically and empirically underdetermined: neither does the hypothesis itself entail the evidential relation between spots and measles, nor does the fact of the spots (see Longino, 1979, 1990, pp. 38–61).

Now, one might think that if there is under determination of theory by evidence, then evidence is ultimately internal to each theory, and alternative theories cannot be compared on the grounds of evidence. Such raises the specter of epistemic relativism and global skepticism, which we will see is a recurrent attack from Longino's critics (see

² While Longino's views on evidence have developed over time, they are a continuation from her landmark (1979) paper "Evidence & Hypothesis" in *Philosophy of Science*, which she acknowledged as the basis for Chapters 2 and 3 of *Science as Social Knowledge* (1990, p. xi).

Sect. 3.1). However, Longino compellingly presented her analysis of under determination as a *solution* to Thomas Kuhn and Paul Feyerabend's problem of insurmountable incommensurability, where the theory-ladenness of observation could leave scientists in different paradigms inescapably speaking past one another. In contrast, Longino preserves the possibility of intersubjectivity for empiricism in the face of theoryladenness: on her view, evidential status is dependent on *relevance*, which is, in turn, dependent on background assumptions. Rather than an irreconcilable difference, disagreement between people over which hypothesis the facts support *might* legitimately require recourse to a discussion that foregrounds, unpacks, and improves those operative but previously implicit assumptions. Because evidential status is contingent on background assumptions, and because hammering out disagreements over "the evidence" often requires social interactions, there cannot be a strict distinction between empirical evidence and social interactions. Whereas Kuhnian-Feyerabendian incommensurability might render evidence irreducibly dependent on the paradigm and thus internal to the theory, Longino's recognition and articulation of the role of background assumptions allows for different evidential assessments to be discussed and overcome through interaction across people with somewhat different background assumptions.

For Longino, background assumptions include contextual values as well as scientific laws, models, logic, and theories, so decisions about evidence based on background assumptions are at times value judgments. Nonetheless, Longino's solution pushes the question of the epistemic justification of theory choice one step back: *why hold certain background assumptions over others?* She writes:

Even though the prospect of an infinite regress prevents one from supposing that the adoption of all beliefs could be evidentially based, there is no *a priori* reason to suppose that there are no criteria at all. *One can ask whether, and if so, which, criteria should determine their acceptance.* (Longino, 1979, p. 55, emphasis added)

We can take the development of Longino's work since 1979, especially *Science as Social Knowledge*, as an elaboration of how to adjudicate background assumptions, empirically or otherwise, with special attention to value-laden background assumptions.

3 Feminist empiricism about values

Among feminist empiricists,³ one increasingly divisive topic is how to be an empiricist about values themselves (Clough, 2020; Solomon, 2012; Yap, 2016). Based on her (1979) analysis of evidential reasoning, Longino (1990) established contextual

³ The term "feminist empiricism" comes from Sandra Harding (1986, p. 24) to designate the feminist epistemology that contends "that sexism and androcentrism are social biases correctable by stricter adherence to the existing methodological norms of scientific inquiry." Harding distinguished this empiricist brand of feminist epistemology that aimed to preserve traditional objectivity from what she called "feminist standpoint theory" and "feminist postmodernism" (1986, pp. 26–27). Though widely influential, Harding's classification of feminist empiricism as an alternative to standpoint theory has been criticized by feminist standpoint advocates (Intemann, 2010; Wylie, 2012), along with her misrepresentation of Longino's position as simply replacing androcentrism in science with value-free, gender-neutral descriptions (Lloyd, 1988a).

empiricism as a new form of *feminist* empiricism: "It is *empiricist* in treating experience as the basis of knowledge claims in the sciences. It is *contextual* in its insistence on the relevance of context—both the context of assumptions that supports reasoning and the social and cultural context that supports scientific inquiry—to the construction of knowledge" (Longino, 1990, p. 219, emphasis added). Yet, several feminist philosophers of science, particularly Sharyn Clough (1998) and Maya Goldenberg (2015), have criticized the way Longino relates empirical evidence to values. This section unpacks both of their critiques, which to varying degrees are based on misrepresentations and misunderstandings of Longino's accounts of empirical evidence and intersubjective objectivity.

We contend that at the heart of the disagreements in this debate are the different epistemological traditions in feminist philosophy of science, which give rise to several competing conceptions of "feminist empiricism" and significant divergence among them over what counts as "empiricism," especially from more radical advocates of holism and "naturalized epistemology" (see Intemann, 2010; Richardson, 2010; Solomon, 2012; Crasnow, 2013; Yap, 2016; Brown, 2020). Longino's contextual empiricism is more akin to Bas van Fraassen's (1980) constructive empiricism; it differs from Clough's (1998) semantic holism from Donald Davidson, who rejects representation-based epistemologies for a more action-based, behavioristic understanding of knowledge that has little space for background beliefs (LePore & McLaughlin, 1985).⁴ Clough's position has been called "feminist *radical* empiricism" for endorsing a "web of *valief*...the all-encompassing network of beliefs and values that is described by feminist empiricists in the Quinean tradition" (Solomon, 2012, p. 435, emphasis added).⁵

It is important to note that the legacy of pragmatism is also at stake, which Clough defines as naturalized, socio-historical, and holistic in the neo-pragmatist tradition of Davidson and Richard Rorty (Clough, 2013b). This contrasts with more classic American pragmatists like John Dewey that are more consistent with Longino's program (see Sect. 3.3; Anderson, 2006; Brown, 2020). Also debated is the social nature of empiricism: For Clough, the empirical bases of epistemology are language and learning, which are inherently social processes; for Longino, one of the reasons that "science is social knowledge" is because background beliefs are learned socially but also because the objectivity of science depends on communal processes of intersubjective criticism, uptake, and response. For Goldenberg (2014, 2015), however, community-based modes of mitigating the harmful effects of values are less practical than more individual, evidence-based assessments—which directly contrasts with Longino's rejection of the dualistic binary of *rational* empirical assessment by individuals vs. *social* negotiation by collectives (Longino, 2002). Given the complex nature of empirical evidence

⁴ Quinean feminist empiricisms also differ those descended from Rudolf Carnap that preserve an a priori (Yap, 2016).

⁵ Sharon Crasnow (2013) places Longino's contextual empiricism solely within feminist empiricism and against feminist standpoint theory. However, Longino aligns herself with practice-based standpoint theory (1990, p. 188), and the contemporary versions of these approaches are compatible under several interpretations (Intemann, 2010; Wylie, 2012). A more significant distinction is with feminist postmod-ernists/posthumanists (e.g., Haraway, 1991, p. 186; Barad, 2007, p. 244), who contrast their approaches with older, more empiricist modes of feminist critique of science (see Richardson, 2010). See footnote 3.

and the dynamic interplay between individuals and communities, it is unsurprising that such disagreements have emerged, even among feminists who share many social and political goals.

Furthermore, much confusion is bred in the discussion by cross-talk about "values," "value judgements," and "background assumptions." Philosophers of science typically use the category *value* to mean an abstract quality that some people take to be good, desirable, or worthy of pursuit, such as simplicity among physicists or sustainability among environmentalists (McMullin, 1982, p. 5; Elliott, 2017, p. 11). Values are thus a normative form of *belief* that guide our *value judgments*, including the appraisal of means and the prizing of ends (Dewey, 1939, p. 5). Clough (1998, 2013a, 2020) focuses on the actual values themselves being evaluated (e.g., anti-Black racism), and Goldenberg (2015) primarily analyzes value judgements (e.g., discounting industry-funded studies), while Longino is interested in the assessment of the background assumptions uberhaupt—including the values, methodologies, laws, models, theories, social norms, and biases—all of which can influence science (see Sect. 2).

In this section, we hope to distinguish legitimate philosophical disagreements from misunderstandings and mischaracterizations. Following our own evaluation of these critiques, this section concludes more positively with an elaboration of how Longino's contextual empiricism explicitly understands values as *heuristics*, which steer the practice of research programs in certain directions that can be empirically fruitful or not for particular uses (Longino, 2008). This sets the stage for our case study in Sect. 4 on the empirical disconfirmation of androcentric values in evolutionary biology.

3.1 Clough's critique: "hasty retreat from evidence" to relativism

We begin with the earliest analysis from Clough (1998), that Longino "retreated" from empirical evidence (which Clough calls "content") toward values and worldviews ("scheme"). Over the past two decades, Clough has advocated for treating values (and value judgments) as having empirical content subject to the gauntlet of human experience (Clough & Loges, 2008; Clough, 2003, 2013a, 2013b, 2020). For instance, she argues that "racist value judgments," like beliefs about racial hierarchies based on biological differences, "express beliefs that are objectively false" in terms of their "cognitive or descriptive content" about the world (Clough & Loges, 2008, p. 77; Clough, 2020). Clough's framework relies on Davidson's critique of "representationalism," referring to the view that theories represent the world by using "language schemes" as a *filter* through which empirical "content" flows (Clough, 1998, p. 91). Accordingly, Clough takes contextual values to be "schemes" in Davidson's sense of passive constraints or filters that screen off content, unaffected by empirical evidence (see also Clough, 2003).

Rather than "the search for better evidence (or content)," Clough criticizes Longino for *instead* promoting "a search for better conceptual filters (or schemes)" (1998, p. 94). This is not to say that Clough takes a naïve positivist view of evidence as unmediated and "simply 'given'," nor that she disagrees with Longino over the contested nature of evidence; their disagreement is instead *how to understand* the interdependency

of values and evidence (pp. 94–95).⁶ Clough claims that, because Longino bases scientific representations on the filtering of personal values and beliefs, she leaves a "metaphysical gap between our theories and the world" that invites "global skepticism" and subjective epistemic "relativism": "All of our representations could be floating free of the world, to varying degrees" (p. 101). Such undesirable consequences, Clough argues, invite the need for a more robustly ontic form of feminist empiricism, focused on understanding the *causal* relations between theories and the world (see Clough, 1998, 2003, 2013b).

While we welcome parts of Clough's positive program, her critique of contextual empiricism is based on three key mischaracterizations of Longino's view. First, the interpretation that Longino abandoned empirical evidence (or content) in favor of values (or scheme) is inaccurate. What Clough perceives as a "retreat" from evidence/content and advance to values/scheme is more charitably understood as an elaboration of *the very category of evidence*. Rather than accepting a simple view of evidence as "the given" or positing an insurmountable incommensurability, Longino's framework allows for a deeper, relational understanding of evidence and disagreements over it based on background assumptions. By making background assumptions transparent through the community's criticism, we can better understand the source of divergence over judgments about evidential relevance—and improve them.

Part of the misunderstanding follows from a mischaracterization of Longino's stance as an ontic view itself, typified by Clough's description of a "metaphysical gap between the subjective end product of belief and the objective external reality the belief is about" (1998, p. 91). Because Clough interprets Longino as a representationalist-someone who distinguishes between internal schemes and external content-she then criticizes contextual empiricism for claims that Longino never actually makes: As Longino has described it, the "gap" is merely a "logical" one about evidence and justification, not a metaphysical gap between mind and matter (1990, p. 58; see Sect. 2). So when Longino talks about this underdetermination gap between data and hypothesis (bridged by background beliefs), she is not talking about an ontic gap between the knower and what is known but an *epistemic* gap between claims (observations and ideas) about the world. This gap between claims is bridged by background beliefs, which are largely incompatible with a standard Davidsonian picture that takes behaviors as the only causally or logically significant factors and ignores intentions, mental representations, and other implicit/unstated beliefs (LePore & McLaughlin, 1985). While a Davidsonian might take aim at Longino's metaphysical quietism, it is uncharitable and incorrect to interpret Longino in such heavily metaphysical terms.

Second, Clough misrepresents Longino's view by falsely claiming that, for Longino, values are merely negative "constraints on reasoning" like Davidson's passive filters (Clough, 1998, p. 96). On the contrary, values for Longino are among the positive epistemic conditions for reasoning that justify background assumptions, in turn *enabling* evidential classifications that structure social interactions about empirical evidence. That is, values can serve as background assumptions and supply these relevance relations that connect data and hypothesis, rendering some data relevant to a given hypothesis and other data irrelevant. Emphasizing their active nature, Longino

⁶ We thank an anonymous reviewer for helping us highlight these agreements between Longino and Clough.

has described values even more dynamically as *heuristics* for inquiry that positively guide knowledge production and the development of theories and models (Longino, 2008; see Sect. 3.3).

Third, and most problematically, Clough (1998, pp. 100–101) mischaracterizes Longino by describing her conceptions of theory choice and evidence as "subjective" and "relativist" because they give up "on the potentially decisive role of evidence." Clough falsely claims that Longino bases theory choice solely on "subjective" values or personal preferences and does not acknowledge "belief-independence of empirical evidence" (1998, p. 106). Clough's scheme/content framing projects this subjective/objective binary onto Longino's framework, inviting her accusation that Longino posits "a split between an inner conceptual world of values and interpretive frameworks and an outer world of unanalyzed data" (p. 94). While Clough agrees with Longino that evidence is often indecisive and that we need to be "examining the historical and political conditions under which data become taken as evidence," she worries that Longino has rendered science into a debate over "interpretative frameworks" (Clough, 2020, pp. 25–26). Instead, Clough contends that the "decision about which values to prioritize itself becomes an investigative project that properly continues at a meta-level using roughly the same kinds of empirical criteria" (2020, pp. 25–26).⁷

In contrast with Clough's framing of Longino approach as "subjective," contextual empiricism maintains that the status of "evidence" is not purely a question of personal interpretation, but rather a relational property between a hypothesis/theory and the data taken to be evidence for it; so evidential status depends partially on background beliefs and partially on data and observations from the empirical world. Here, the distinction in Longinian thought between *evidence* and *data* is significant: it is precisely during the transformation of multipotent data/observations of the world into relevant evidence via background beliefs that values can play their most influential roles. That mediation of data and hypotheses by background assumptions is where scientific criticism ought to be aimed, which Clough appears to misunderstand by taking Longino to have a subjectivist view of evidence based on her own objectivist alternative. Empirical evidence is relational and *intersubjective* (neither wholly external nor wholly internal), and as we discuss next, Longino instead takes values, while privately held, to be socially accessible and sometimes even testable.

3.2 Goldenberg's critique: "values as evidence," rather than negotiated

While Clough (1998) argues that Longino has "retreated from evidence" and invited skepticism, Goldenberg (2014, 2015) makes a different critique, more about scientific objectivity: she argues that Longino's norms for interactive objectivity are not the only (or the better) way to adjudicate values within science. Along with Clough, she advocates a "values as evidence" approach that accepts theoretic underdetermination of evidence but also promotes more direct empirical evaluation of values than what she sees in Longino. Yet, unlike Clough, Goldenberg focuses on what she calls Longino's "feminist criterion for inclusive community arbitration of the values that inextricably

 $[\]overline{}^{7}$ We thank an anonymous reviewer for helping us show these overlaps between Longino and Clough and their disagreements.

enter into scientific reasoning" (2015, p. 4), contending that they undermine more practicable and rational forms of empirical assessment:

That [underdetermination] gap is filled by the contextual values that mediate evidential relations. *Those values, notably, 'may not be subject to empirical confirmation or disconfirmation'* (Longino, 1990, p. 75). Background assumptions [according to Longino] *are thereby not subject* to the same empirically driven modes of scrutiny that scientific reasoning affords to data. (Goldenberg, 2015, p. 12, emphasis added)

Note how Goldenberg moves from Longino's hedged claim that values "may not be subject" to empirical test to the much stronger claim that values "are thereby not subject" to such tests.

Goldenberg focuses on entrenched problems in medicine involving the hegemony of the pharmaceutical industry, conflicts of interest, and lack of transparency—all of which Longino would argue reduce interactive objectivity. Yet, Goldenberg contends that biomedicine is nonetheless structured well enough for individual clinicians and researchers to deliver treatments with good standards of evidence. She considers a hypothetical doctor, who is trying to decide whether to offer her patient a new therapy with seemingly promising results (Goldenberg, 2015, p. 23). The doctor soon realizes that the positive findings in the trial she initially consulted were, in fact, likely the result of industry-funding bias, using low dosages of the standard therapy for the controls to get a (lucrative) false positive. Confirmed by the additional knowledge that industry funding predicts pro-industry conclusions, the doctor comes to the "*reasonable* decision within the confines of the information available to her" to not offer her patient the therapy (Goldenberg, 2015, p. 23, emphasis added).

According to Goldenberg, this thought experiment illustrates how clinicians, in the face of a seemingly less than objective research community, can gather empirical evidence (an underpowered control and the industry's "tricks of trade") to support value judgements (ignore this study's positive results) based on value-laden assumptions (distrust industry studies), which are themselves empirically evidenced. She concludes that clinicians can use their judgment based on empirically confirmed values *rather than* "the cumbersome recourse to epistemic communities" a la Longinian objectivity (Goldenberg, 2015, p. 24). Unlike this more direct use of "values as evidence," she charges that Longino's proposed social solutions fail because value-laden "background assumptions are thereby not subject to the same empirically driven modes of scrutiny that scientific reasoning affords to data" (2015, p. 12). Because of the effectiveness of individually using empirically evidenced value judgments to combat systemic financial conflicts of interest throughout the community, Goldenberg concludes that Longino's social approach is "impractical and unnecessary" (2014, 2015, p. 5; contrast with Goldenberg, 2021).⁸

⁸ Interestingly, in Goldenberg's (2021) monograph on vaccine hesitancy, she expresses much more agreement with Longino's overall approach than in these earlier papers from 2014 and 2015. Most notably, she advocates for the importance of transformational criticism for objectivity (2021, p. 97) defined as "intersubjective agreement" (p. 196, n. 11) rooted in both everyday social dynamics (p. 122) and the diversity within scientific communities (p. 152). We thank an anonymous reviewer for pointing out this more recent convergence.

While it would be problematic if Longino shielded values from direct empirical scrutiny, this is not an accurate representation of Longino's position. Goldenberg's reading of Longino is challenged by a discussion near the end of *Science as Social Knowledge*, in which Longino writes:

Some background assumptions may involve *conceptual, metaphysical, and normative dimensions* that elude assessment by strict empirical criteria. *Others may be subject to fairly straightforward empirical assessment.* Arguments that use factual hypotheses to undermine or support claims about values provide good subjects for study. (Longino, 1990, p. 183, emphasis added)

She then considers the studies on boys' and girls' mathematic abilities. Longino argues that the traditional, gender-neutral assumption that test preparation is uniform for boys and girls is "deficient on straightforward empirical grounds" citing empirical studies evidencing sexist practices, such as different treatment by teachers that promote different skills (p. 183). Thus, she contends that these other lines of empirical evidence undermine the value-laden assumption rooted in the conservative value of gender neutrality in the classroom (i.e., that there is no differential gender-based discrimination in math classrooms) that supports business-as-usual education. However, other operative background assumptions (e.g., there is only one form of mathematical ability) have not been "investigated systematically" (p. 183). Without empirical studies, we are left with conventional intuitions about cognitive abilities and normative critiques of them, such that "the interests of [already recognized individuals of mathematical ability] are served by not challenging the assumptions" (p. 184).

Therefore, contra Goldenberg, Longino did not claim that empirical assessment of values (or background assumptions) was impossible outright and in all cases, but rather that testability depends on whether relevant studies are available and whether the problem is currently approachable as an empirical one. Fleshing out the assumptions in need of empirical support through conceptual criticism may enable testing and confirmation in the future (see Sect. 4).

Digging a little deeper, we see how Goldenberg's critique is based on four mischaracterizations of Longino's norms for objectivity. First and foremost, she relies on a false dichotomy between *empirical* assessment (testing and confirmation) and *social* assessment (criticism and negotiation). Goldenberg claims that the only way a contextual empiricist can adjudicate between different values is by non-empirical methods, namely "under scrutiny by the democratic, inclusive, and responsive community of knowers" (2015, p. 14). She contrasts Longino's social approach with her (individualist) one, exemplified by the doctor sitting alone in her office, where such values are arbitrated "using many of the same empirical modes of inquiry used to scrutinize empirical claims" (p. 15).

Yet, the dichotomy between individual empirical assessment and social processes of deliberation is precisely the dualism Longino undermined with her analysis of empirical evidence (see Sect. 2). Social and evidential assessment are not alternatives but instead *inseparable* during the process of science: social interaction improves one's understanding of the evidence for/against a hypothesis by drawing out the underlying assumptions. Furthermore, the *goodness* of evidential reasoning relies on public community standards, such as what counts as a good explanation, good statistical practices,

and good enough evidence to support a theory (Longino, 1990). While Goldenberg (2014, 2015) seems to disagree over the crucial importance of social processes (contrast with Goldenberg, 2021), Clough (2013b) might be in more agreement here with Longino on the necessarily social nature of epistemology.

Goldenberg's reliance on the empirical-social dichotomy for methods of assessment is clear in her own thought experiment, which she takes to exemplify the self-sufficiency of the former without dependence on the latter. Accordingly, she downplays the importance of shared community standards for assessing evidence (Longino, 1990, p. 77) and, thus, potentially overstates the decisiveness of evidence for individual value judgments (like discounting results from industry-funded studies) in politicized contexts. These community standards include peer-review processes like double-masking and statistical standards accepted across the profession, including the infamous 5% standard level of significance (Gigerenzer, 2004; Porter, 1996). Further related to Goldenberg's example are requirements to disclose funding sources and register all trials—standards that have arisen because of community concerns about conflicts of interest (see Holman & Elliott, 2018).

The second major issue is how Goldenberg (2015) mischaracterizes the operation of Longinian norms as "arbitration by democratic vote" (p. 26) competing with the rationality of individual judgment and thus lacking "rational content" (p. 24). Longino's shift to the social level for accessing, applying, or challenging norms does not preclude the possibility of the rationality of individuals' judgments conforming to social norms and processes, as Goldenberg implies, unless one accepts the two as strictly dichotomous. Even if we grant (as Longino would) that only sometimes can values be assessed empirically, this could not be done in the isolated, individualistic manner suggested by Goldenberg. Many judgments would be derivative of community standards for good science, and others require social criticism to unpack their evidential relations. Thus, while we might grant that individuals can devise "what appears to be a reasonable decision" after the fact (Goldenberg, 2015, p. 23), its reasonableness would nonetheless depend on community practices. By artificially segmenting the social and individual aspects of objectivity, Goldenberg suggests that scientists and clinicians can and should evaluate evidence in isolation from social processes. Such heavy emphasis on individual judgment is badly misleading-especially in science where collaboration, trust, and group deliberation are essential (Andersen & Wagenknecht, 2013; Wilholt, 2013; Wray, 2014).

Third, and relatedly, Goldenberg ignores the social process of science by looking only at the products and outcomes available for expert judgment. She claims that.

There was *no recourse to a social process* of critical scrutiny required either in order to justify the values invoked or the conclusion that the physician drew. Instead, the relevant contextual values rested on empirical claims that were legitimately arbitrated using the same modes of scientific reasoning to which all empirical evidentiary claims can be subjected. (Goldenberg, 2015, p. 24, emphasis added)

While framed as *either* social *or* empirical, one should ask: where is this clinician getting her knowledge about the influence of Big Pharma? Why does she trust those

sources instead of industry's response? Whence come her standards and critical attitude?

Much of Goldenberg's (2015, p. 24) concern appears to be the impracticality and inconvenience of "cumbersome recourse to epistemic communities" ostensibly from Longino. Yet, this is an overly literal reading of contextual empiricism, as if the clinician must wait to make her decision while "the jury is out." The clinician can appeal to public standards and depend on community-wide practices of peer-review while still having an empirically grounded decision. Moreover, the cited studies about industry bias have in fact resulted from interactions between researchers and their critics, often from outside medicine, precisely because of the hegemony of medicalization and entrenchment of commercial interests. For example, because of widespread biasing practices in the research on using selective serotonin reuptake inhibitors (SSRIs) to treat depression—such as not publishing negative studies, ghostwriting, and not reporting suicidal behavior-it took external criticisms from lawyers and journalists to document harmful effects (Jukola, 2015). Part of the reason one can trust this research about industry bias is because of the rigorous criticism it has undergone, such as the meta-analyses evidencing how widespread the problem is, especially in medicine (Bekelman et al., 2003; Lundh et al., 2018).

The final and most important problem with Goldenberg's argument is that it appears to exonerate the need for community-level reforms in medicine by focusing on individual-level judgment as a sufficient solution: "No appeal to an idealized community of knowers needs to be made to come to a reasonable decision, and so the great effort required to build this epistemic community need not discourage health-care workers and researchers from pursuing smaller scale remedies to the problems of evidence-based healthcare" (2015, p. 26). Because of the interdependence of individual and social processes, we contend that a two-level approach would prove more effective than either alone.

Moreover, we worry that this shift toward the individual for rational deliberation unduly acquits the biomedical community of much needed structural change. While some individual practitioners might be able to recognize funding bias, many others will not on their own. Furthermore, industry can bias research communities without corrupting any individual researcher. For instance, the pharmaceutical industry created a consensus around the efficacy of (ineffective) antiarrhythmic drugs using the strategy of selectively funding researchers with pro-industry conclusions—leaving the less favorable ones underfunded—and it can use similar tactics effectively whenever there is methodological diversity within a merit-based system (Holman & Bruner, 2017). The composition of research groups is subject to the biasing pressures levied by the pharmaceutical industry, so it would be fruitless to put our faith in individuals to simply "follow the evidence" when they lack funds and other support. Instead, the solution involves building social networks of trust and heavily reducing researchers' financial dependency on industry (Holman & Bruner, 2015; Holman & Elliott, 2018; Wilholt, 2013).

Thus, Goldenberg's critique downplays the interaction of social and empirical assessment in science, falsely suggesting the two are alternatives rather than parts of an integrated process. Longino did not argue for a "recourse to epistemic communities." Instead, she *contextualized* how all individual judgment is necessarily dependent on

community standards. Furthermore, Longino recognized that some value-laden background assumptions "may be subject to fairly straightforward empirical assessment," depending on the state of the science; even where there is a lack of reliable empirical data, conceptual and normative assessments would be appropriate for value judgements *until such relevant empirical data could be produced* (1990, p. 183). In either case, fleshing out and improving empirical reasoning depends in part on community standards and social processes.

3.3 Contextual empiricism revisited: values as heuristics

If it is too simplistic to think of values always as evidence per se, then what would be a better understanding? This section elaborates our alternative account of values as heuristics, in which values *can* be empirically assessed in a more partial sense of *communal validation* in particular contexts of specific uses according to their *fruit-fulness*, i.e., their empirical, historical, social, and explanatory success in a variety of contexts.⁹ Here, we supplement the contextual-empiricist approach with Deweyian pragmatism for understanding the normative relations between values and empirical evidence. Just like with models in scientific practice, direct validation and complete confirmation of values is not possible (Oreskes et al., 1994).

In her more recent work, Longino describes values as having "heuristic but not probative power," meaning that they can offer reliable guidance but not mathematical proof, usually transmitted to scientists as part of their "background" during scientific apprenticeship, and their learning of the history of their discipline through stories of success and failure (Longino, 2008, p. 74). She illustrates this historical evolution of scientific heuristics with examples, such as the traditional reliance in medical research on the value of *simplicity* that guided researchers to select primarily white cisgender men as "the norm." While this heuristic facilitated many advancements in medical trials throughout much of the twentieth century, it also resulted in the systematic failure to understand the safety and effectiveness of treatments in people of color, cisgender women, and gender non-conforming people. Today, the heuristic of diversity and inclusion has arisen in response to these empirical failings, which were also equally political and ethical failures. This newer diversity-orienting heuristic has guided us toward a more clinically useful, more empirically accurate, and more inclusive set of practices for medical trials, such as representative sampling and sub-population analysis (see Epstein, 2007).

To develop this account of values as heuristics further, we point others toward the insights of pragmatist philosophers of science Elizabeth Anderson (2004, 2006) and Matthew Brown (2020), who build on ideas from Dewey's theory of inquiry and valuation (e.g., 1939). Anderson emphasizes the instrumental value of values and

⁹ Historically, many philosophers of science have taken the scientific value of fruitfulness to be *merely pragmatic*, as opposed to its epistemic strength (e.g., McMullin, 1982). For example, Kuhn (1977, p. 322) claimed that the value of fruitfulness lies in the reputational and professional prospects that a successful theory promises for a scientist's career. This traditional pragmatic-but-not-epistemic framing is precisely what we oppose. Instead, we contend that fruitfulness offers *jointly* pragmatic and epistemic value because empirical successes include every-day experiential fruits as well as more scientific and societal yields (see Anderson, 2004).

their dynamic nature in the process of inquiry: "Dewey argued that value judgments function as *tools* for uncovering data for better living...We *test* our value judgments by living in accordance with them, and seeing whether we find the results satisfactory" (Anderson, 2006, p. 4, emphasis added). Such "tests" are humanistic "experiments in living," with the successes of feminism contrasting women's lived dissatisfaction with sexist values (Anderson, 2006, p. 5).

Through a case study of feminist research on divorce, Anderson (2004) illustrates how values can function like empirical hypotheses, able to be confirmed or disconfirmed empirically by experience and observation if held non-dogmatically. In the 1990s, a team of social scientists led by Abigail Stewart uncovered new facts on the benefits of divorce that were unwelcome by the dominant community of researchers. Those conservative researchers held the following heuristics as background: they endorsed "traditional family values," assuming that divorce "breaks up" a (heterosexual) family and necessarily harms children. Thus, looking for only the harms of divorce—and ignoring its benefits—the traditionalists were unable to collect a fair sample of evidence because they did not believe that positive evidence could exist. Accordingly, their hypotheses were so biased that the traditionalists only confirmed the values with which they began, amassing evidence that divorce harmed women, children, and families. In contrast, Stewart and her feminist-scientist collaborators documented the *mixed* nature of divorce, including its potential benefits as well (e.g., offering spouses an opportunity for growth), thus generating a more complete set of evidence. Thus, Anderson argues that when values act *fallibly* (opening possibilities rather than foreclosing them), they positively guide hypotheses toward empirically fruitful research.

One way to judge whether a value as a heuristic is successful or not is by its empirical fruits: does it help to produce well-specified models that are well-supported by observations, experiments, and other knowledge rooted in lived experience? In Anderson's divorce case, the (non-feminist) traditionalist heuristic led to limited research questions, e.g., "What are the *costs* of divorce for men, women, and their families?", so the responsive models that the traditionalists developed were only partial. The traditionalists did not ask or learn about any of the *benefits* of divorce because benefits were not imagined as possible or responsive answers to their originating question and framework. This contrasts with the feminists' more ambivalent and open-ended research questions, e.g., "What are the *consequences* of divorce for men, women, and their families?" The question here shows its superiority by its broader empirical range of possible and responsive answers. The feminist value was a more fruitful guide toward understanding the empirical phenomenon of divorce—thus showcasing the direct scientific yields of this heuristic in social science on families.

Building on Anderson's ideas and Dewey's logic of inquiry, Brown (2020) argues that empirical inquiry is always *situational*, rather than radically holist in the manner advocated by Clough. All empirical inquiry, including science, begins with an indeterminate situation and ends with a judgment, and what falls in between is highly contingent on the context: "it is not their form or essence that suits [values] to be evidence, but rather...their ability to play the *functional* roles of evidence that suits them to be evidence" (Brown, 2020, p. 99, emphasis added). Nonetheless, Brown rightly

notes that it would be an "equivocation across functional distinctions" to say that values are equivalent to evidence, in the radically holist manner advocated by Clough and Goldenberg, because evidence is a functional relation of empirical support between data and theory, while values are normative grounds for practical judgments in situations of uncertainty (Brown, 2020, p. 206).

When understood more broadly as heuristics, we see better how values are not usually evidence per se (though they might at times play justificatory roles of empirical support for/against a hypothesis), but rather guides for *motivating* empirical inquiry, *ordering* data classification, and *directing* theory building. Here, Brown suggests that values play an even more significant role in empirical inquiry that *precedes* evidence, for the empirical support we gain is brought about by the values that guided the search in their direction (Brown, 2020, p. 100). That is, values provide frameworks for asking questions, constructing models, and defining relevant evidence, thus providing the epistemic setup for any empirical data to become evidence.

Recalling Anderson's case study, feminist researchers approached the topic of divorce with an open mind *because of their feminist values*, considering both the positive possibility that divorce could liberate women from abusive husbands and the negative potential that it might also enable negligent men to leave their wives and increase their own fortunes in the process. We might call this mode *interested inquiry*, the way in which values as heuristics direct us toward evidence for liberation. This is perhaps an even wider, more social sense of fruitfulness than philosophers of science typically have, under which values guide knowledge collection about aspects of living that matter to us.

Thus, rather than treating values as evidence per se, we suggest it is often better to think of values as *heuristic tools*, active and responsive to particular uses with resistance or ease, signaling failure or success for the task at hand. Such pragmatic interplay between values and evidence echoes Longino's own words: "what constitutes 'our world' is not a given but a product of the interaction between the external material reality that is 'the world' and our own pragmatic and intellectual needs" (1990, p. 221). Now, one might object that, if contextual empiricists advocate for understanding values as heuristics, then maybe Longino really did "retreat from evidence" as Clough claims. That is, a heuristic seems like a merely subjective, theoretical "scheme" that filters objective, empirical "content," so Longino's empiricism would still be at risk of skepticism and relativism. But this would be a misunderstanding of our approach: a heuristic is an active framework—at least partially subject to community-wide empirical evaluation—held by a community for building models that answer their research questions, not merely a passive set of personal beliefs or idiosyncratic schemes. In the next section, we use a case study to show how, contrary to this subjectivist view of heuristics, their use is actually subject to communal validation according to their fruitfulness, resulting in a partial, indirect dis/confirmation within certain contexts on empirical grounds.

4 A case study of contextual empiricism: disconfirming sexist values in biology

The stakes of these debates about empirically assessing values include the continuing threat of oppressive ideologies in science. What can Longino offer us against unfair biases and in favor of feminism and social justice more generally? Goldenberg and others charge that Longino's brand of feminist empiricism lacks the "empiricist grounds for endorsing feminist values over, say, androcentric and sexist values in feminist research" (2014, p. 26). One might worry that "the feminism in [Longino's] gap feminist empiricism is [merely] a contingent feature of recent successful science criticism" rather than something essential to its feminist values themselves (Solomon, 2012, p. 439). Likewise, Clough suggests that we need to see the problems associated with sexist and racist values as an "empirical failure" rather than simply an ethical one (2013a, p. 74).

While Longino's view might *legitimate* feminist values in science prima facie, Anderson contends that "it does not help us evaluate the different ways that values might be deployed in inquiry" (Anderson, 2004, p. 2; see also Intemann, 2005, 2010). Without support for feminist values per se, Longino might be *too inclusive* and *permissive* regarding sexist, racist, and other oppressive values in science. Dan Hicks (2011, p. 337) calls this "the Nazi Problem": that "Longino's account of objectivity requires the active cultivation of historically excluded and marginalized groups," extended to white supremacists, anti-Semites, misogynists, and homophobes.

In this section, we maintain that contextual empiricism can in fact supply empirical grounds for excluding (delegitimizing) oppressive values like sexism from science as empirically disconfirmed. To begin, recall that Longino advocates for a "practice-based" feminism that frames "science as *practice* rather than content, *process* rather than product" (1990, p. 188, emphasis added). Understanding Longino's feminism as practice-based and process-oriented, we can better see how contextual empiricists are committed to feminism *procedurally*, particularly through shared communal norms. She contends that feminist scientists practice "oppositional science" that challenges dominant assumptions that exclude marginalized groups (Longino, 1990, p. 214). Likewise, to promote diversity and inclusiveness in the process, more objective communities practice democracy by attempting to mitigate hierarchies of authority among scientists (1990, p. 78, 2002, p. 128).

Now, there is genuine disagreement here between Longino and her critics over these procedural features in her account because of her focus on a diversity of *values* (and background assumptions) rather than other kinds of diversity, such as a diversity of social positions and embodied experiences (see Intemann, 2010). Yet, these procedural norms exemplify anti-oppressive commitments, and so they provide practitioners the grounds to challenge unequal power relations within science, even if their effectiveness and inclusiveness is limited. From our perspective, such criticism minimizes the additional normative force supplied by other Longinian norms of objectivity, especially shared public community standards like standards of evidence (1990, p. 77; 2002, p. 130). As Longino writes, "constitutive values provide a check on the role of contextual values and cultural assumptions. These constraints include *empirical* and *conceptual* evaluation of assumptions" (1990, p. 223, emphasis added).

This section uses the case of androcentrism in biological research on the evolution of female orgasm to demonstrate how a contextual empiricist like Elisabeth Lloyd (2005) can use community standards to critique the empirical, epistemic strength of sexist values, in terms of effects on the logic of research questions (see also Lloyd, 1993, 2013, 2015, 2021). Here, evidence has empirically undermined widespread background assumptions that prioritize cisgender men and male bodies and conflate cisgender women's sexuality with reproductive capacity. Echoing Longino, we maintain that while it is *not always* possible, there are some situations when oppressive values like androcentrism and heteronormativity are disconfirmed on fairly straightforward empirical grounds for specific uses.¹⁰ This case study shows that Longino's contextual empiricism, contrary to the claims of her critics, is indeed empiricist about values, particularly when it comes to the empirical fruits of feminist heuristics in fields like sexology, genetics, and anatomy.

4.1 Empirically assessing values in the case of the female orgasm

In her investigation of 21 evolutionary explanations of female orgasm, Lloyd applies and further develops Longino's contextual empiricism (2005, p. 220). Lloyd demonstrates that certain values played major roles in causing the empirical deficiency of explanatory models for the evolution of female orgasm. She focuses on Longino's requirement that objective science comes from communities with publicly recognized standards for evaluating theories, hypotheses, and observations. Lloyd also uses Anderson's emphasis on the partiality of evidence to ground her contextual empiricist analysis, appealing "to all facets of available empirical data," especially "data that... are inconsistent with one's assumptions" (2005, pp. 244–245).

Lloyd treats the self-styled "ardent adaptationists," that is, those extreme advocates of natural selection who are strongly committed to finding adaptive, functional accounts for each and every trait, while assuming that it arises from natural selection (Alcock, 1987). She investigates their adherence to a distinct set of standards of evidence compared to the main body of the evolutionary community, which considers a broader range of evolutionary factors, including genetic linkage, phyletic inertia, developmental byproduct, as well as adaptations (further elaborated and supported in Lloyd, 2015, 2021). Lloyd describes contrasting roles for the scientific value of adaptationism with more contextual background assumptions, focusing on three specific sets (2005, pp. 233–36):

- Androcentrism: looking at the world from an exclusively male perspective, neglecting a unique female point of view;
- 2. *Human uniqueness*: emphasizing the differences between humans and our recent common ancestors and relatives; and
- 3. *Heteronormative procreative focus*: assuming that all evolutionarily significant sex is heterosexual intercourse between males and females with the direct potential for offspring.

¹⁰ For more ethics/justice-oriented approaches for criticizing sexist science, see Kourany (2010) and Hicks (2014).

Showing how these value-laden background assumptions influence empirical reasoning and practice, she weaves a complex web of interactions of values, heuristics, and evidence. Accordingly, Lloyd emphasizes Longino's requirement that scientific communities engage in critiques of their own background assumptions using publicly recognized standards: "we can see that the [above three assumptions] are implicated in partial treatments of the data, in which relevant data are ignored. The same is true for adaptationism" (2005, p. 248).

Following Longino, Lloyd evaluates the background assumptions of androcentrism, adaptationism, and heteronormativity relative to empirical data and communal norms (Lloyd, 2005, p. 249). In an early articulation of the logic and pragmatics of research questions (Lloyd, 2015, 2021), she spells out her expansion of Longino's community standards, including:

- 1. Standards of what questions to ask;
- Standards of what empirical evidence is relevant and appropriately established in answering those questions; and
- 3. Standards of what kind of explanation is appropriate, or what answer to the question is suitable (Lloyd, 2005, p. 250).

She then compares the various sub-communities' uses of these sets of standards, emphasizing empirical evidence.

What Lloyd found in the totality of evidence for the 21 available evolutionary theories for female orgasm was very disappointing. Nearly all of the various "pair bond" theories included an empirical assumption well-known to be false among sexologists—as well as many cisgender women—namely the assumption that whenever a heterosexual couple engages in vaginal intercourse, the woman reliably has orgasm (90–100%). In fact, the most recent frequency for unassisted intercourse without manual clitoral stimulation is 21–30% (Shirazi et al., 2018), with other studies giving even lower estimates (from 4 to 18%) of cisgender women who reliably orgasm with vaginal penetration alone (see Mahar et al., 2020).

In addition, the two accounts using the then-trending "sperm competition" theory also failed on the evidence, since the ten features associated with sperm competition were all missing in human cisgender men, and the statistics of the single experimental study were extremely flawed (Lloyd, 2005, pp. 179–219; Dixson, 2012). Moreover, all 20 of the adaptation-based explanations required a firm positive correlation between genetic fitness (number of offspring) and orgasm rate, which was undermined empirically after the publication of her 2005 book, showing orgasm rate has no effect on number of offspring (Zietsch & Santtila, 2013).

Amongst the shambles, Lloyd found that only one theory of the evolution of female orgasm's maintenance in the population was positively supported by a number of pieces of empirical evidence. The "indirect selection" theory, founded in stabilizing selection in males, says that the female orgasm reflex evolved along with the clitoris and other structures, features, and properties, because those features were evolutionarily advantageous in male members of the species. Because these features enabling the reflex developed early in the process of emerging from embryo to infant to adult, both the male and female fetus (and infant and adult) develop them. Lloyd (2005, 2015) calls this the "byproduct/bonus" theory, as female orgasm in primates, including humans

assigned female at birth, is the result of the stabilizing selection on male primates, which affords contemporary cisgender women and other humans born with a clitoris the benefit of peaks of sexual pleasure.

The anthropologist and sociobiologist Donald Symons (1979) created this byproduct theory, carefully eliminating competing theories and advancing empirical evidence for this explanation. Unfortunately for the reader, his well-supported indirect selection theory was buried in such an unfortunately androcentric book that, naturally, some feminists took objection to his work (see Lloyd, 2005, pp. 139–140). These critics, both philosophers and biologists, objected that Symons' theory of female orgasm did not take account of the full complexity of female sexuality, that it devalued women's sexual experiences, and that it still treated female orgasm androcentrically as a proxy for male orgasm (see Wakil, 2021).

Disregarding the indirect selection theory's origin in this clearly male-centered book, Lloyd concluded that it was nevertheless the most empirically well-supported approach of all 21 theories, and that because of the strength of the evidence, it should be considered seriously and more research done (2005, pp. 220–257). She further maintained that this byproduct/bonus explanation was less biased with androcentric values than the others because it separates women's sexuality from its reproductive function (Lloyd, 2005, p. 238). Furthermore, simply because this "bonus" is causally derivative from a male trait does not entail the female trait has any less value (since natural selection is not the arbiter of human values). On the contrary, androcentric heuristics have obscured many of the relevant observations *supporting* the by-product hypothesis.

As Lloyd herself has argued (2005, p. 220), her treatment of this case is exemplary of the Longinian framework: she shows where evidential reasoning depends on value-laden background assumptions like androcentrism and uses external (feminist) criticism to improve the empirical adequacy of the explanations.¹¹ Critics of our linking Lloyd and Longino might contrast the apparent singularity of Lloyd's conclusion with Longino's permissiveness: that is, they might object that while Longino is an empiricist, she supports a thorough-going "theoretical pluralism," in which different approaches with different background assumptions "constitute a nonunifiable plurality of partial knowledges" (Longino, 2006, p. 127; see also 1990, p. 230). For instance, her more recent work has contrasted competing scientific approaches to studying human behavior (quantitative behavioral genetics vs. developmental systems theory vs. socioenvironmental approaches, etc.). Here, Longino famously argued there is "a plurality of approaches generating accounts of the etiology of individual behavioral dispositions that are not reducible to some fundamental level of causation, not integratable into a single comprehensive account, and not empirically commensurable in a way that would permit elimination of rivals in favor of one" (2013, p. 135, emphasis added). Accordingly, one might reason that contrary to Lloyd's analysis in her 2005 book, a true contextual empiricist should embrace not just one theory but several theories, as a plurality is necessary for understanding complex phenomena like female orgasm.

¹¹ Longino herself agrees with the authors that Lloyd's work fits well within the framework of contextual empiricism (pers. comm., 2019).

This objection is misleading because it mistakes pluralism for a commitment to a plurality of theories, when in fact pluralism is more of an *attitude* that contrasts with the dominant monism among scientists who assume they will eventually eliminate plurality: "Monism exists primarily as a default assumption underlying polemical and philosophical arguments about these [various scientific] approaches [to behavior]. Pluralism is best understood as an attitude to adopt with respect to the multiplicity of approaches in contemporary sciences" (Longino, 2013, p. 138). Thus, a pluralist of Longino's stripe would be satisfied with holding incompatible theories but remain open to the possibility that, in some cases, there may be only one theory with adequate empirical support.¹² Accordingly, a pluralist should not support the remaining 20 theories (in the literature at the time) after Lloyd and others have shown that they were contradicted by abundant relevant evidence, along with their constitutively value-laden background assumptions.

Thus, while one might see this as a possible tension between Lloyd and Longino, it actually exposes a shared response to the threat of relativism. Lloyd's analysis illustrates how Longino's contextual empiricism is capable of avoiding a pernicious relativism and the "global skepticism" that "all of our representations could be floating free of the world, to varying degrees" (Clough, 1998, p. 101; see also Intemann, 2005): The pluralism of contextual empiricism here is better understood as rooted in fallibilism (the refusal to accept any belief as certain) rather than radical global skepticism (the active doubting of all beliefs in the absence of absolute certainty). Advocating for a plurality of approaches, in line with Longino and Lloyd, is rooted in the historically established ability of science to fail and the tight alliance between dominant science and hegemonic powers. Lloyd herself notes that female orgasm can be described in a variety of ways, including "relatively reductionist biological descriptions" involving muscle spasms and blood flow in the pelvic and genital area, or a more psychologically robust description involving hormones and neurotransmitters since "female orgasm turns out to be quite a bit more neurologically complicated than the simple knee-kick reflex" (2005, p. 23). In typical fallibilist manner, Lloyd emphasizes in her introduction that her analysis is partial and provisional: "Though at this time I find no credible evidence that female orgasm is an adaptation, I am open to such a finding. Female orgasm may very well turn out to be an adaptation, exquisitely designed for some special but obscure function. None of my arguments is meant to rule this possibility out" (2005, p. 17). The key here is that because of the joint commitment to fallibilism and pluralism, contextual empiricism can accommodate both the parsimony that Lloyd (2005) promotes with her work on female orgasm as well as the permissiveness that Longino (2013) has advocated in her work on human behavior (for more on pluralism in evolution, see Lloyd, 2001).

In line with such pluralism, one might also support a complementary theory on female orgasm. Inspired by Lloyd's research, Pavličev and Wagner (2016) have pursued another related research question about evolution and orgasm involving its deepest origins among mammals. They offered empirical evidence supporting a view that the orgasmic reflex derived from the reflex contraction of the ovary upon release of an ovum (Pavličev et al., 2019). This hypothesis (concerning adaptive spasms of the ovary

¹² We thank Helen Longino for her help in wording this response.

in the deep past) may seem to undermine the byproduct/bonus hypothesis because it appeals to an archaic adaptation, yet, as Lloyd has argued, the two explanations are compatible¹³: An ovarian spasm is not functionally *equivalent* to a modern spasm of the circle of muscles and tissues around the vaginal opening (a common contemporary physiological definition of a female orgasm), nor do these authors claim that it is. And Pavličev and Wagner (2016) are clear that they do *not* see female orgasm as being adapted by evolution to its present state in humans (see also Pavličev et al., 2019).

We contend that these successes of Lloyd's case study support Longino's idea that empirically assessing values like androcentrism is a communal process. However, one might worry that while we have argued that shared norms provided the resources to criticize sexist science, the value of androcentrism itself was one of evolutionary biology's entrenched communal norms. That is, one might see Lloyd's intervention as simply another "traditional logical empiricist framework" that is continuous with "business as usual" science, premised on the naïve idea that "gender bias—in common with all biases—produces inferior science, sometimes called 'bad science'" (Solomon, 2012, pp. 436–438). Thus, it could appear that the success achieved by Lloyd's intervention was not because of Longino's "transformative criticism" based on specific values like feminist heuristics, but rather because of the simple self-correcting mechanisms of biased science based on empirically false assumptions.

In response, we contend that Lloyd has shown how her analysis of evidence and values is a natural extension of the Longinian picture: Contrary to Solomon's (2012, p. 437) characterization of Lloyd (2005) as a logical empiricist "trained to try to keep scientific reasoning free of all bias," Lloyd's actual framework of analysis, the Logic and Pragmatics of Research Questions (Lloyd, 2015; Morrison, 2021), emphasizes the importance of values and the community work entailed in science. It is the active scientific community of researchers who determine not only the research questions being pursued, but also their possible and responsive answers, and what the standards of acceptable evidence might be in any given case (Lloyd, 2015). Contextual empiricism offers a social and evidential analysis of the relations between values, questions, answers, and their empirical support. The relevant standards and values derive from the community and are applied by individual scientists and smaller groups (Lloyd, 2021). Empirical evidence does in fact work against androcentrism and ardent adaptationism, in accordance with the empiricist values held in the wider evolutionary community-a vivid and concrete example of the Longinian picture. But what does this communal process of disconfirming values actually look like under Longino's view?

4.2 The process of confirming/disconfirming values empirically

Lloyd's (2005) case study demonstrates how there is a constellation of values working as heuristics and interacting synergistically with evolutionary and scientific norms of explanation to generate community accepted science. Here, adaptationism, heteronormativity, and androcentrism have worked in concert, supplementing the (misperceived)

¹³ Lloyd's talk in October 2016, "Evaluating the New Theory of the evolution of female orgasm," Humanities Center, Yale University.

empirical strength of functionalist accounts and their supporting evidence. The heuristics of androcentrism and heteronormativity were *reinforced* by ardent methodological adaptationism, only to be detected and undermined as such by a feminist, scientific, and philosophical critic (see Lloyd, 2013).

Thus, there is no one, single value or heuristic being confirmed or disconfirmed in relation to community norms. We cannot get a unique determination (akin to Popperian falsification) in science; however, we are not left with radical underdetermination or subjectivism in a given case like ours. The values embedded in various evolutionary biological accounts are based on community norms of explanation and confirmation of evolutionary models, which follow certain standards, explored independently and confirmed by their empirical fruits (Lloyd, 1988b).

As Longino noted, at times there may be no studies available to empirically test a given value-laden background assumption: these can involve conceptual and normative dimensions at first before empirical data become available (1990, p. 183). As one reviewer of Lloyd's book wrote, tongue-in-cheek, "The sad fact is that, for now, all statements about the evolution of the female orgasm are conjectures in an empirical vacuum. To advance the debate, we need data... In short, it's time to collect data. Without it, the debate will remain like sex sometimes is: furious, empty and anticlimactic" (Judson, 2005, pp. 916–17). While there was significant evidence against adaptationist accounts at that time, conceptual critiques by Lloyd directly inspired further empirical studies (e.g., Wallen & Lloyd, 2011, p. 780; Zietsch & Santtila, 2011, p. 1097, 2013, p. 253; Shirazi et al., 2018, p. 606; Blair et al., 2018, p. 2), eventually building the capacity to more directly empirically evaluate these heuristics.¹⁴ For example, the androcentrism that had anchored early adaptationist accounts was decisively challenged by later studies demonstrating the anatomical reason for why some cisgender women have orgasm reliably with (unassisted) vaginal intercourse while most do not (Oakley et al., 2014; Shirazi et al., 2018; Vaccaro, 2015; Wallen & Lloyd, 2011).

Because of new interest in the byproduct explanation from Lloyd (2005), new empirical studies became available confirming and discussing this hypothesis, which had been largely neglected for decades, and more recent scientific reviews now take this non-adaptation hypothesis as an equal contender in the field (e.g., Welling, 2014). Adaptationist explanations *assumed* a strong fitness effect of female orgasm, but geneticists inspired by Lloyd's critique found no correlation with genetic fitness, once they investigated this assumption directly through two twin studies (Zietsch & Santtila, 2011, 2013). These geneticists acknowledged their debt to Lloyd's critique explicitly: "The heat of this debate has recently intensified ...after the 2005 publication of a provocative book by Elizabeth [sic] Lloyd" (Zietsch & Santtila, 2011, p. 1097). This genetic evidence constitutes an existential challenge to all adaptationist theories, which now needed to explain why, if selection produced female orgasm, there is no evidence of adaptedness.

All this shows the importance of social processes of transformative criticism for investigating value-laden background assumptions empirically: In this instance, scientists set out to test the byproduct approach to determine empirical support for it.

¹⁴ All of these empirical studies cite Lloyd (2005) as foundational. For instance, Talia Shirazi and her colleagues note the major significance of Lloyd's critique for the framing of their research (2018, p. 606).

Lloyd's critical feminist work had large impacts on positive scientific developments, especially in genetics, anatomy, and sexology, changing the very research questions of these fields. Empirical evidence is *cultivated* in the context of inquiry guided by values operating as heuristics including background assumptions, dissatisfaction with the current state of the evidence, and desires for new evidence. *Developing such empirical evidence requires values and can disconfirm values as more or less fruitful heuristics for building models to make sense of the empirical world.* Thus, contra Clough's (1998) claim that contextual empirical "content," we see here how feminism can operate as an *active* heuristic that *positively* influences the empirical bases with which we can assess value-laden background assumptions.

Lloyd's book has continued to motivate genetic, physiological, and behavioral research expanding knowledge about female orgasm. In sexology, these include the first statistically significant findings on the orgasm rate for bisexual and lesbian women (Garcia et al., 2014). This study, led by Justin Garcia including Lloyd herself, highlighted the issue of women's neglected sexual satisfaction: while heterosexual men orgasm 86% of the time during sex with a familiar partner, for heterosexual women the rate is 24% less (62%) (Garcia et al., 2014). Given that less a women had higher rates of orgasm (75%), one likely explanation-pursued by feminists many decades earlier (e.g., Hite, 1976)—is that the difference results less from female anatomy and more from poor technique, gender roles, and sexual attitudes. In the time since Garcia et al. (2014) established the sizable gap between lesbian and heterosexual women's orgasm frequencies, there has been an explosion of research on gendered "orgasm gaps" (see review by Mahar et al., 2020), including empirical studies on the most effective behaviors for stimulating partnered female orgasm across heterosexual, bisexual, and lesbian women (Blair et al., 2018; Frederick et al., 2018) as well as study of the gender differences in sex questionnaires regarding cisgender women's orgasm with intercourse (Shirazi et al., 2018).¹⁵

Now that we have seen how feminist values have motivated this research on disparities in orgasm frequency by gender and sexuality, one might ask: *why* have feminist values steered researchers in this particular direction? Recalling the idea of interested inquiry (see Sect. 3.3), feminist heuristics guide us toward creating a world that promotes sexual pleasure for all genders and sexualities and, thus, a world free from sexual guilt, shame, and frustration. Many cisgender women have been made to feel shame at not having orgasm from unassisted heterosexual intercourse; yet, through feminist guided-research, we can now show empirically that variation in women's orgasm frequency is partially explained by *anatomy* and *cultural attitudes*: these include the distance between the clitoris and the vagina and the likelihood of penile stimulation with penetration (Vaccaro, 2015; Wallen & Lloyd, 2011), as well as androcentric, heteronormative attitudes toward sex that prioritize vagina-around-penis/penis-in-vagina intercourse over more direct clitoral stimulation (Mahar et al., 2020).

Thus, at the root of these sexual inequalities are harmful gender roles that neglect women's pleasure, normalized by the dated Freudian theory that clitoral orgasms are

¹⁵ Frederick et al. also explicitly note the scientific impact of Lloyd's book for their field (2018, p. 274). See footnote 14.

"immature" and that becoming a woman involves developing the ability to orgasm with (unassisted) heterosexual vaginal penetration; yet, those people born with bodies assigned female at birth and socialized as cisgender women have long experienced the inadequacy of this account of "frigidity" and have developed feminist critiques of it rooted in the current science (e.g., Beauvoir, 1953; Koedt, 1970). Their experience of such sexual inequality and neglect has enabled scientists to generate knowledge about the existence of the phenomena of orgasms gaps because of their scientific desire to create models and collect relevant data to quantify it more systematically (see Lloyd, 2005, pp. 26–27; Shirazi et al., 2018). Values are rooted in our embodied experiences of the world, and feminist values operate as guides for creating new scientific facts, particularly knowledge in service of liberation for all.

5 Not values as direct empirical evidence, but values as heuristic tools in the logic of research questions for specific studies

Therefore, with this case it is clear that value-laden background assumptions operating as *heuristics* are empirically assessable. From the perspective of contextual empiricism, a value in the very abstract sense (like androcentrism or feminism) is disconfirmed only limitedly and indirectly (see also Solomon, 2012). However, when a value functions as a "heuristic" that *constructs* evolutionary models, *motivates* scientific data collection, and *prompts* biological explanations about female orgasm, it is more strongly dis/confirmed within a specific domain like primatology, human anatomy, and human genetics. As contextual empiricists, we judge these values by their empirical fruits in our scientific pursuits.

We have defended contextual empiricism's approach to assessing values empirically as a more viable form of feminist empiricism than the direct "values as evidence" approach developed in opposition to a caricature of Longino. In fact, the very language of "values as evidence" coined by Goldenberg (2015) comes not from older feminist studies of values in science like Longino's work, but rather from more recent work in the *environmentalist* line of philosophy of science, specifically Heather Douglas's groundbreaking (2009) book. There, Douglas introduces the conception of direct vs. indirect influences of values, which crucially prohibits scientists' use of values as evidential support for their acceptance/rejection of hypotheses (2009, p. 97). That is, values could either influence scientific decisions *directly*, functioning as the primary warrant or reason in support of the judgment, or *indirectly*, influencing the standards of evidence that a scientist requires for accepting a claim. While Douglas accepts the direct function for non-technical judgments, like choosing research topics, she is critical of values playing a direct role throughout the technical parts of the scientific process: "Values should never suppress evidence, or cause the outright rejection (or acceptance) of a view regardless of evidence" (2009, p. 113). It seems to us that feminist empiricists like Goldenberg have appropriated Douglas's categories (values as evidence vs. values as standards of evidence) and her normative position (when the direct/indirect role is appropriate) and then misattributed them to Longino.

Given that Clough and Goldenberg do in fact wish to support a more direct approach to treating values as evidence—in which there is a more radical holism and a "facticity of values" in general (Goldenberg, 2021, p. 57)—this connection with Douglas helps contrast our middle-ground approach: While contextual empiricism does not go quite as far as Douglas to completely prohibit direct use of values (in the more technical contexts), our understanding of values as heuristics means that they are less directly evaluated than Clough and Goldenberg suggest is typically possible, as they are embedded in the background of a community in the logic of research questions. Values as heuristics can be subject to empirical disconfirmation as we have shown, but more indirectly through the failure of their use in generating useful research questions, motivating more complete data collection, and building empirically adequate models.

In sum, we have elaborated Longino's vision of values as heuristics, and we have shown the applicability of this empiricist account to values in biology. We expect this robust, contextualist view to have many successful applications in other domains of science as well. Over three decades after its publication, we might have more to learn from what we have forgotten about *Science as Social Knowledge* than what we think we remember.

Acknowledgements We gave an earlier version of this paper at the 2022 International Conference on Engaging Ethics and Epistemology in Science at Leibnizhaus in Hannover, Germany, and received many helpful comments from the audience. We have benefited from the generous feedback of many colleagues and friends, including Kevin Elliott, Ryan Ketcham, Anna Leuschner, Helen Longino, Naomi Oreskes, Ryan O'Loughlin, and Manuela Fernández Pinto, as well as three anonymous reviewers.

Funding Open access funding provided by SCELC, Statewide California Electronic Library Consortium.

Declarations

Competing interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Alcock, J. (1987). Ardent adaptationism. Natural History, 96(4), 4.

- Alcoff, L. M. (2006). Commentary on Elizabeth Anderson's "uses of value judgments in science." MIT Symposia on Gender, Race and Philosophy, 2(1), 1–4.
- Andersen, H., & Wagenknecht, S. (2013). Epistemic dependence in interdisciplinary groups. Synthese, 190(11), 1881–1898.

- Anderson, E. (2004). Uses of value judgments in science: A general argument, with lessons from a case study of feminist research on divorce. *Hypatia*, 19(1), 1–24.
- Anderson, E. (2006). Replies to my critics: Reply to Alcoff, Clough, Janack, and Mills. MIT Symposia on Gender, Race and Philosophy, 2(1), 1–6.
- Barad, K. (2007). Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning. Duke University Press.
- Beauvoir, S. (1953). The second sex. Alfred A. Knopf.

Bekelman, J. E., Li, Y., & Gross, C. P. (2003). Scope and impact of financial conflicts of interest in biomedical research: A systematic review. *Journal of the American Medical Association*, 289(4), 454–465.

Blair, K. L., Cappell, J., & Pukall, C. F. (2018). Not all orgasms were created equal: Differences in frequency and satisfaction of orgasm experiences by sexual activity in same-sex versus mixed-sex relationships. *The Journal of Sex Research*, 55(6), 719–733. https://doi.org/10.1080/00224499.2017.1303437

Bleier, R. (1984). Science and gender: A critique of biology and its theories on women. Pergamon Press.

- Brown, M. (2020). Science and moral imagination: A new ideal for values in science. University of Pittsburgh Press.
- Clough, S. (1998). A hasty retreat from evidence: The recalcitrance of relativism in feminist epistemology. *Hypatia*, 13(4), 88–111.
- Clough, S. (2003). Beyond epistemology: A pragmatist approach to feminist science studies. Rowman & Littlefield.
- Clough, S. (2013a). Feminist theories of evidence and biomedical research communities: A reply to Goldenberg. Social Epistemology Review and Reply Collective, 2(12), 72–76.
- Clough, S. (2013b). Pragmatism and embodiment as resources for feminist interventions in science. Contemporary Pragmatism, 10(2), 121–134.
- Clough, S. (2020). Using values as evidence when there's evidence for your values. *Philosophy in the Contemporary World*, 26(1), 5–37. https://doi.org/10.5840/pcw2020261/21
- Clough, S., & Loges, W. E. (2008). Racist value judgments as objectively false beliefs: A philosophical and social-psychological analysis. *Journal of Social Philosophy*, 39(1), 77–95. https://doi.org/10.1111/j. 1467-9833.2007.00412.x
- Collins, P. H. (1986). Learning from the outsider within: The sociological significance of Black feminist thought. Social Problems, 33(6), 14–32.
- Crasnow, S. (2013). Feminist philosophy of science: Values and objectivity. *Philosophy Compass*, 8(4), 413–423.
- Dewey, J. (1939). Theory of valuation. International Encyclopedia of Unified Science, 2(4), 1–66.
- Dixson, A. (2012). Primate sexuality. 2nd ed. Oxford University Press.
- Daston, L., & Galison, P. (2007). Objectivity. Zone Books.
- Douglas, H. (2009). Science, policy, and the value-free ideal. University of Pittsburgh Press.
- Elliott, K. (2017). A tapestry of values: An introduction to values in science. Oxford University Press.

Epstein, S. (2007). Inclusion: The politics of difference in medical research. University of Chicago Press.

Frederick, D. A., St. John, H. K., Garcia, J. R., & Lloyd, E. A. (2018). Differences in orgasm frequency among gay, lesbian, bisexual, and heterosexual men and women in a US national sample. *Archives of Sexual Behavior*, 47(1), 273–288.

Garcia, J. R., Lloyd, E. A., Wallen, K., & Fisher, H. E. (2014). Variation in orgasm occurrence by sexual orientation in a sample of US singles. *The Journal of Sexual Medicine*, 11(11), 2645–2652.

- Gigerenzer, G. (2004). Mindless statistics. The Journal of Socio-Economics, 33(5), 587-606.
- Goldenberg, M. J. (2014). Diversity in epistemic communities: A response to Clough. Social Epistemology Review and Reply Collective, 3(5), 25–30.
- Goldenberg, M. J. (2015). How can feminist theories of evidence assist clinical reasoning and decisionmaking? Social Epistemology, 29(1), 3–30.
- Goldenberg, M. J. (2021). Vaccine hesitancy: Public trust, expertise, and the war on science. University of Pittsburgh Press.
- Haraway, D. J. (1991). Simians, cyborgs, and women: The reinvention of nature. Routledge.
- Harding, S. (1986). The science question in feminism. Cornell University Press.
- Hicks, D. J. (2011). Is Longino's conception of objectivity feminist? Hypatia, 26(2), 333-351.
- Hicks, D. J. (2014). A new direction for science and values. Synthese, 191(14), 3271–3295.
- Hite, S. (1976). The Hite report: A nationwide study of female sexuality. Seven Stories Press.
- Holman, B., & Bruner, J. (2015). The problem of intransigently biased agents. *Philosophy of Science*, 82(5), 956–968.

- Holman, B., & Bruner, J. (2017). Experimentation by industrial selection. *Philosophy of Science*, 84(5), 1008–1019.
- Holman, B., & Elliott, K. C. (2018). The promise and perils of industry-funded science. *Philosophy Compass*, 13(11), e12544. https://doi.org/10.1111/phc3.12544
- Holman, B., & Wilholt, T. (2022). The new demarcation problem. Studies in History and Philosophy of Science, 91, 211–220. https://doi.org/10.1016/j.shpsa.2021.11.011
- Hubbard, R. (1979). Have only men evolved? In R. Hubbard, M. S. Henifin, & B. Fried (Eds.), Women look at biology looking at women. Schenkman Publishing Co.
- Intemann, K. (2005). Feminism, underdetermination, and values in science. *Philosophy of Science*, 72(5), 1001–1012.
- Intemann, K. (2010). 25 years of feminist empiricism and standpoint theory: Where are we now? *Hypatia*, 25(4), 778–796.
- Intemann, K. (2015). Distinguishing between legitimate and illegitimate values in climate modeling. European Journal for Philosophy of Science, 5(2), 217–232.
- Jasanoff, S. (1990). The fifth branch: Science advisers as policymakers. Harvard University Press.
- Judson, O. P. (2005). Anticlimax. Nature, 436(7053). 916-917. https://doi.org/10.1038/436916a
- Jukola, S. (2015). Longino's theory of objectivity and commercialized research. In S. Wagenknecht, N. Nersessian, & H. Andersen (Eds.), *Empirical philosophy of science: Introducing qualitative methods into philosophy of science* (pp. 127–143). Springer International Publishing.
- Koedt, A. (1970). *The myth of the vaginal orgasm.* New England Free Press. Retrieved from https://docs. wixstatic.com/ugd/63d11a_f1aa3818f1b6471aa113c28aeb5130f7.pdf
- Kourany, J. A. (2010). Philosophy of science after feminism. Oxford University Press.
- Kuhn, T. S. (1977). Objectivity, value judgment, and theory choice. In *The essential tension*, (pp. 320–339). University Of Chicago Press.
- LePore, E., & McLaughlin, B. (Eds.). (1985). Actions and events: Perspectives on the philosophy of Donald Davidson. Blackwell.
- Lloyd, E. A. (1988a). Review: The science question in feminism by Sandra Harding. Isis, 79(2), 308-309.
- Lloyd, E. A. (1988b). The structure and confirmation of evolutionary theory. Greenwood Press.
- Lloyd, E. A. (1993). Pre-theoretical assumptions in evolutionary explanations of female sexuality. *Philosophical Studies*, 69, 139–153.
- Lloyd, E. A. (1995). Objectivity and the double standard for feminist epistemologies. *Synthese*, 104(3), 351–381.
- Lloyd, E. A. (2001). Units and levels of selection: An anatomy of the units of selection debates. In R. S. Singh (Ed.), *Thinking about evolution: Historical, philosophical, and political perspectives* (pp. 267–291). Cambridge University Press.
- Lloyd, E. A. (2005). The case of the female orgasm: Bias in the science of evolution. Harvard University Press.
- Lloyd, E. A. (2013). The evolution of female orgasm: New evidence and response to feminist critiques. In F. de Sousa & G. Munevar (Eds.), *Sex, reproduction, and Darwinism* (pp. 109–125). Pickering and Chatto.
- Lloyd, E. A. (2015). Adaptationism and the logic of research questions: How to think clearly about evolutionary causes. *Biological Theory*, 10(4), 343–362.
- Lloyd, E. (2021). Adaptation: Elements in the philosophy of biology. Cambridge University Press.
- Lloyd, E. A., & Schweizer, V. J. (2014). Objectivity and a comparison of methodological scenario approaches for climate change research. *Synthese*, 191(10), 2049–2088.
- Longino, H. E. (1979). Evidence and hypothesis: An analysis of evidential relations. *Philosophy of Science*, 46(1), 35–56.
- Longino, H. E. (1990). Science as social knowledge. Princeton University Press.
- Longino, H. E. (2002). The fate of knowledge. Princeton University Press.
- Longino, H. E. (2006). Theoretical pluralism and the scientific study of behavior. In S. H. Kellert, H. E. Longino, & C. K. Waters (Eds.), *Scientific pluralism* (pp. 102–131). University of Minnesota Press.
- Longino, H. E. (2008). Values, heuristics, and the politics of knowledge. In M. Carrier, D. Howard, & J. Kourany (Eds.), *The challenge of the social and the pressure of practice: Science and values revisited* (pp. 68–86). University of Pittsburgh Press.
- Longino, H. E. (2013). Studying human behavior: How scientists investigate aggression and sexuality. University of Chicago Press.

- Lundh, A., Lexchin, J., Mintzes, B., Schroll, J. B., & Bero, L. (2018). Industry sponsorship and research outcome: Systematic review with meta-analysis. *Intensive Care Medicine*, 44, 1603–1612.
- Mahar, E. A., Mintz, L. B., & Akers, B. M. (2020). Orgasm equality: Scientific findings and societal implications. *Current Sexual Health Reports*, 12(1), 24–32. https://doi.org/10.1007/s11930-020-00237-9
- Marchetti, G., & Marchetti, S. (Eds.). (2016). Facts and values: The ethics and metaphysics of normativity. Routledge.
- McMullin, E. (1982). Values in science. In P. D. Asquith & T. Nickles (Eds.), Proceedings of the 1982 Biennial meeting of the PSA (Vol. 2, pp. 3–28). Philosophy of Science Association.
- Morrison, M. (2021). The models are alright: A socio-epistemic theory of the landscape of climate model development (no. 28315089) [Doctoral Dissertation, Indiana University, Bloomington]. ProQuest Dissertations Publishing.
- Oakley, S. H., Vaccaro, C. M., Crisp, C. C., Estanol, M. V., Fellner, A. N., Kleeman, S. D., & Pauls, R. N. (2014). Clitoral size and location in relation to sexual function using pelvic MRI. *The Journal of Sexual Medicine*, 11(4), 1013–1022.
- Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263(5147), 641–646. https://doi.org/10.1126/science.263. 5147.641
- Pavličev, M., & Wagner, G. P. (2016). The evolutionary origin of female orgasm. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 326(6), 326–337.
- Pavličev, M., Zupan, A. M., Barry, A., Walters, S., Milano, K. M., Kliman, H. J., & Wagner, G. P. (2019). An experimental test of the ovulatory homolog model of female orgasm. *Proceedings of the National Academy of Sciences*, 116(41), 20267–20273.
- Porter, T. (1996). Trust in numbers: The pursuit of objectivity in science and public life. Princeton University Press.
- Richardson, S. (2010). Feminist philosophy of science: History, contributions, and challenges. Synthese, 177(3), 337–362.
- Schiebinger, L. (2001). Has feminism changed science? Harvard University Press.
- Shirazi, T., Renfro, K. J., Lloyd, E. A., & Wallen, K. (2018). Women's experience of orgasm during intercourse: Question semantics affect women's reports and men's estimates of orgasm occurrence. *Archives of Sexual Behavior*, 47(3), 605–613.
- Solomon, M. (2012). The web of valief: An assessment of feminist radical empiricism. In S. L. Crasnow & A. M. Superson (Eds.), *Out from the shadows: Analytical feminist contributions to traditional philosophy* (pp. 435–450). Oxford University Press.
- Symons, D. (1979). The evolution of human sexuality. Oxford University Press.
- Vaccaro, C. M. (2015). The use of magnetic resonance imaging for studying female sexual function: A review. *Clinical Anatomy*, 28(3), 324–330.
- Van Fraassen, B. C. (1980). The scientific image. Oxford University Press.
- Wakil, S. (2021). Objectivity and orgasm: The perils of imprecise definitions. Synthese, 199(1), 2315–2333. https://doi.org/10.1007/s11229-020-02886-8
- Wallen, K., & Lloyd, E. A. (2011). Female sexual arousal: Genital anatomy and orgasm in intercourse. *Hormones and Behavior*, 59(5), 780–792.
- Welling, L. L. M. (2014). Female orgasm. In V. A. Weekes-Shackelford & T. K. Shackelford (Eds.), Evolutionary perspectives on human sexual psychology and behavior (pp. 223–241). Springer.
- Wilholt, T. (2013). Epistemic trust in science. The British Journal for the Philosophy of Science, 64(2), 233–253.
- Wray, K. B. (2014). Collaborative research, deliberation, and innovation. *Episteme*, 11(3), 291–303.
- Wylie, A. (2012). Feminist philosophy of science: Standpoint matters. Proceeding and Addresses of the American Philosophical Association, 86(2), 47–73.
- Yap, A. (2016). Feminist radical empiricism, values, and evidence. *Hypatia*, 31(1), 58–73. https://doi.org/ 10.1111/hypa.12221
- Zietsch, B. P., & Santtila, P. (2011). Genetic analysis of orgasmic function in twins and siblings does not support the by-product theory of female orgasm. *Animal Behaviour*, 82(5), 1097–1101.
- Zietsch, B. P., & Santtila, P. (2013). No direct relationship between human female orgasm rate and number of offspring. *Animal Behaviour*, 86(2), 253–255.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.