Philosophy of Science (2026), **00**, 1–9
doi:10.1017/psa.2026.10232

PSA PHILOSOPHY OF SCIENCE ASSOCIATION
PHILOSOPHY OF SCIENCE

SYMPOSIA PAPER

Bespoke Knowledge

Bennett Holman

Associate Professor of History and Philosophy of Science Underwood International College Yonsei University, Korea, Republic of
Email: bholman@yonsei.ac.kr

(Received 15 April 2025 UTC; revised 31 May 2026 UTC; accepted 01 June 2026 UTC)

Abstract

A central finding of social epistemology is that phenomena desirable at the individual level might have hidden, undesirable effects that emerge on the social level. These emergent effects have important implications for the management of values in science. Using a historical case study of the safety of hormonal pregnancy tests, I argue that the meta-analysis of studies risks compounding value influence in undesirable ways. In statistical analysis, researchers must either constrain values such that they do not affect the measure of central tendency or accept that they have created "bespoke knowledge," knowledge that cannot be reused or exported to other contexts.

1. Introduction

In "The New Demarcation Problem," Bennett Holman and Torsten Wilholt (2022) argue that the original advocates of the value-free ideal of scientific inquiry cited three reasons. Specifically, they claimed that preventing nonepistemic values from influencing scientific reasoning helps create science that is veracious, authoritative, and universal. Accordingly, any rationally guided replacement for the value-free ideal must either serve these same ends or be accompanied by an argument for why these ends should not be served. Although little contemporary attention has been paid to universality, the current philosophical discourse is clear about the other dangers. Significant violations of veracity (i.e., departures from search for the truth) lead to "wishful thinking," where values lead scientists to conjure confirmation for something they merely wish to be true (Anderson 2004; Brown 2013; de Melo-Martín and Intemann 2016). When values cause significant violations of authority (i.e., knowledge that is not trusted and trustworthy), it leads to coercive paternalism and violations of individual autonomy. The concern here is that values can be pernicious when scientists' idiosyncratic values are foisted upon the public (Alexandorva 2017; Betz 2013; John 2019). But what of violations of universality?

Perhaps they are not of concern. Indeed, Wendy Wagner (2022) has argued that solving the new demarcation problem (i.e., demarcating legitimate value influences

© The Author(s), 2026. Published by Cambridge University Press on behalf of Philosophy of Science Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/psa.2026.10232 Published online by Cambridge University Press

from illegitimate ones) is composed of two separate problems. The first is ensuring that values are not used in ways that create ends-oriented biases (i.e., science that lacks veracity). The second occurs when the values scientists employ diverge from those held by the public (i.e., science that lacks authority). More recently, Inmaculada de Melo-Martín (2024) has argued that these are the only two concerns relevant to solving the new demarcation problem, at least insofar as all other genuine concerns can be reduced to solving these two issues.[1] In this article, I will show that universality (i.e., producing knowledge that can be used for purposes not intended by the researchers) is a distinct scientific good.

In section 2, I first present a case study of the use of hormone pregnancy tests, which situates the kind of methodological choice I wish to examine. In section 3, I consider whether the standard argument from inductive risk could have been used to support this choice. This question will be answered in section 4. It will turn out that the answer is a qualified no, and the reason that the argument from inductive risk cannot be applied illuminates the problem to which I wish to draw our attention. In section 5, I formalize these concerns into a trilemma. In section 6, I conclude the argument by showing that problems arising from nonuniversality do not reduce to concerns about either veracity or authority.

2. Methodological decisions and meta-analysis

In the late 2010s, a scientific dispute emerged over the safety of a drug that had been withdrawn from the British market 40 years prior. There is much that could be said about the events that led to the withdrawal of the drug (Olszynko-Gryn et al. 2018), but what I wish to focus on is the contemporary dispute. Specifically, in assessing the strength of the evidence, present-day researchers split over the acceptability of methodological choices in the original studies and whether they invalidated the contemporary use of meta-analysis to form an all-things-considered judgment of the historical evidence. Because the current dispute turns on details of the original studies, we must first revisit the research that ultimately precipitated the withdrawal.

Prior to the introduction of the stick tests, which have now become a ubiquitous means of detecting a new pregnancy, the methods for early detection were substantially different. In the 1950s, the best available option was to provide one's general practitioner with a urine sample, which would be sent to a laboratory and injected into a frog to see if she laid eggs (a positive indication of pregnancy). In 1958, a new option became available in which a doctor could prescribe a woman a large dose of estrogen. If she were not pregnant, the estrogen would induce menstruation. The ability for a woman to determine whether she was pregnant privately was especially attractive prior to the legalization of abortion. Use of the hormone pregnancy test (HPT) became widespread (Olszynko-Gryn et al. 2018).

The first cloud of suspicion arose when a study published in *Nature* suggested that HPTs might be harmful to the developing fetus (Gal et al. 1967). Despite this and other similar studies, HPTs remained on the market until 1978, when a large study organized by the British government confirmed the association (Greenberg et al.

[1] Both Wagner (2022) and, especially, de Melo-Martín (2024) provide excellent summaries of the concerns expressed regarding how the inclusion of values might lead to violations of veracity and authority. For a response, see Holman and Branch (unpublished).

https://doi.org/10.1017/psa.2026.10232 Published online by Cambridge University Press

1977). However, although the evidence was ultimately strong enough to spur regulators to remove HPTs from the market, when families ostensibly harmed by HPTs attempted to sue the manufacturers of HPTs, the evidence was judged to be insufficient to demonstrate causation.

For decades, it seemed that the story would end here. But in the 2010s, the declassification of regulatory documents, as well as the discovery of internal documents from the manufacturers, reignited a patient rights group's campaign (Olszynko-Gryn et al. 2018). In 2017, the efforts of patients gained national news coverage and prompted the Commission on Human Medicine to form an expert working group (EWG) to reexamine the evidence. The result was not what patient advocates had hoped for. The official report reaffirmed the conclusion that the evidence did not support a causal association between the use of HPTs and congenital malformations (UK Commission on Human Medicines 2017). Dissatisfied with this result, the patients' rights group reached out to the Oxford Centre for Evidence-Based Medicine and asked it to review the evidence. The team, led by the center's director, Carl Heneghan, conducted a meta-analysis of the data and found, contrary to the EWG report, that children whose mothers took an HPT were significantly more likely to have a congenital malformation (Heneghan et al. 2019).[2]

Both the EWG and the Oxford team looked at the same body of evidence, yet they reached different conclusions. Although there were many small disputes, the primary reason for the divergence stemmed from whether it was appropriate to conduct a meta-analysis or whether, as maintained by the EWG, combining the studies provided a misleading result. In support of the latter contention, the EWG raised three issues that threaten the reliability of the Heneghan et al. meta-analysis: recall bias, channeling bias, and confounding (UK Commission on Human Medicines 2019). I will here focus on recall bias and whether the standard argument from inductive risk could have been relied on by Heneghan et al. as a response to the methodological concerns raised by the EWG.

The central issue arises because a significant portion of the original studies were based on retrospective interviews that asked mothers whether they had taken an HPT. The interviews occurred after the child was born and thus (generally speaking) after the mother knew whether or not the child had a congenital condition. The EWG's concern stems from the fact that "using normal babies as a control potentially increases the risk of recall bias as mothers of children with a malformation may be more likely to remember different exposures" (UK Commission on Human Medicines 2019, 18). In support of this concern, the review cites Nora and Nora (1975), who noted that "considerable probing" was often required to uncover the fact that women had taken an HPT.

Because we will dwell on the details, we need to get precise about exactly how this concern is playing out in the data and the analysis thereof. The EWG (presumably) grants that the types of interview techniques used by researchers did help some mothers correctly remember taking an HPT. The concern is that this more thorough questioning would interact with the mother's knowledge of whether her child was born with a congenital anomaly. For example, if women whose child had a congenital malformation were psychologically motivated to find some external cause to blame,

[2] In the interest of transparency, it should be noted that I was a part of this team.

https://doi.org/10.1017/psa.2026.10232 Published online by Cambridge University Press

they might be more likely to recall taking an HPT than women who had healthy babies. If this methodology did engender a recall bias among women whose children had a congenital malformation, this would increase the risk of a false positive (i.e., of declaring HPTs a cause of congenital malformations when they are actually safe).

This brings us to the central point I wish to consider: Could Heneghan et al. (2019) simply justify these methodological decisions by providing a standard argument from inductive risk? After all, even if the EWG is right that thorough questioning introduces a bias, it seems quite clear that whether to use it or not involves a value-laden trade-off. Suppose that the alternative to thorough questioning incorrectly classifies a number of women as not having taken an HPT, but it is equally likely to do so for both women who had healthy babies and women whose children had a congenital abnormality. In this case, the comparison is not biased by the collection method, but because it is less accurate, it has a lower power to detect a difference if it actually exists. In this scenario, there would be an increased risk of a false negative (i.e., declaring that HPTs are safe, when they are actually hazardous).

3. Methodological decisions and the argument from inductive risk

According to the argument from inductive risk, because the researchers face a trade-off between false negatives and false positives, they have an obligation as moral agents to rely on values to make the methodological choice that best balances the trade-off (Douglas, 2009). For example, we should require more evidence that a vaccine is safe before releasing it to the public than we would require to conclude that a belt buckle reliably holds up pants. The societal cost of releasing the vaccine on the market if it is unsafe or ineffective is far greater than a few sagging jeans, so we should demand more extensive testing of the vaccine.

Douglas (2009) takes this point and then observes that if we accept that values can rightly influence our choice of an evidential threshold, then by the same line of reasoning, scientists should also bring their values to bear on selecting a methodology, characterizing the evidence, and interpreting the data.[3] Because choices here also generate trade-offs between false positives and false negatives, values should play a role in these decisions too.

Put somewhat differently, consider the methodological decision involved in HPT research of whether or not to engage in thorough questioning of women in ascertaining whether they had taken an HPT. I have already indicated that thorough questioning of women makes false negatives less likely and false positives more likely

[3] Note that many of these choices will create (and will be expected to create) a bias in the results. To take a classic example from Douglas (2009), a researcher has to decide how to classify an ambiguous rat liver slide regarding whether it shows signs of cancer. Douglas argues that values can influence this decision, for example, choosing to classify it as cancerous because it better aligns with a precautionary set of values. To see that it creates a bias, suppose that, from a purely veratistic perspective, the slide should be classified as cancerous. In such a case, the values align with the data interpretation, but they are not influencing it. Accordingly, if values do influence such choices, then they do so when veratistic considerations would dictate one decision (perhaps only weakly), but value considerations tip the scale and lead the researcher to make a different decision. In such a case, the decision will take the results away from the result the researcher thinks is most accurate, which is to say, the researcher thinks that it creates a bias. The researcher also thinks that this bias is justified by their values (they are "erring on the side of caution").

https://doi.org/10.1017/psa.2026.10232 Published online by Cambridge University Press

(relative to the alternative). But this is not quite true, or better to say, this is true assuming that everything else stays the same. In particular, it assumes that the evidential threshold (i.e., the *p*-value of the statistical analysis) stays the same. But consider another methodological alternative: Instead of engaging in thorough questioning, HPT researchers could have adjusted their *p*-value to achieve a different balance of false positives and false negatives.[4]

Again, Douglas (2009) observes that methodological choices can have effects equivalent to changing our evidential thresholds. Accordingly, if we accept that values are permissible in determining our evidential thresholds, then values should be acceptable in determining which methodological choice we make. This line of reasoning has been called the "downstream conjecture" (Elliott and Richards 2017, 263). Accordingly, if one rejects the value-free ideal, it might seem that Heneghan et al. (2019) can justify the methodological decision of the original HPT researchers (e.g., Nora and Nora 1975) by asserting that this decision correctly balances false negatives and false positives. But as I will show in the following section, this option is not available.

4. The drawback of tailor-made science

To see why Heneghan et al. (2019) cannot justify previous methodological decisions by claiming they properly manage inductive risk, it will be crucial to make a clear distinction between the original studies conducted by various researchers and the meta-analysis conducted by Heneghan et al. It will turn out that if inductive risk is properly managed at the level of the individual study, it will be poorly managed in the meta-analysis. In this section, I will review Stephen John's (2019) "value-apt ideal" and then return to our case study to show why Heneghan et al. cannot defend the methodological choices in the underlying studies by appeal to an argument from inductive risk.

John's (2019) value-apt ideal is an ostensible replacement for the value-free ideal. John argues that when a scientist asserts some claim, they are simultaneously conveying to their audience that the listener has a good reason to accept the claim as true. However, if the claim is dependent on values that the scientist has but the audience does not, then the audience member may not have good reason to believe the claim. Indeed, asserting a value-dependent claim while obscuring the value dependence violates the autonomy of the audience members. Notice, however, that this line of reasoning does not imply that scientists should never use values in their scientific reasoning, only that they should make sure that the values they use align with those of the recipient of the knowledge claim. Science need not be value-free; it need only be value-apt.[5]

[4] To get an intuitive sense of this, note that the ratio of false positives to false negatives can be made arbitrarily high or low by adjusting the evidential thresholds. If no evidence is enough to declare an effect, then there cannot be a false-positive finding (because you never conclude that there is an effect). If any evidence is enough, then there can never be a false negative (because you will always conclude there is an effect). Obviously, both extremes are absurd standards, but somewhere in between these alternatives are the ratios of both methodological choices. Some combination of one methodology and some evidential standard would produce the same balance of false negatives to false positives as the other methodology with a standard evidential threshold (e.g., $p < .05$).

[5] A similar idea has been put forward Irzik and Kurtulmus (2021) with their notion of "enhanced epistemic trust."

https://doi.org/10.1017/psa.2026.10232 Published online by Cambridge University Press

So, let's consider the happiest of scenarios in which the same set of values is shared by the researchers of the initial studies, the EWG, Heneghan et al., and indeed, all of the relevant users of this research. Further suppose that the EWG is correct that recall bias is a problem in the original studies, but also that the shared values of all involved are such that the thorough questioning of mothers involved in the study leads to the correct balance of false negatives and false positives in the original studies. In such a case, the original research will be value-apt, but the meta-analysis will not be.

This is because if there is a bias in the data, then as the sample size increases, the risk of a false negative decreases, and the risk of a false positive increases.[6] If these risks change, then a person whose values justified the trade-offs when the sample size was small must judge the trade-off as undesirable when the sample size becomes sufficiently large. Functionally speaking, a meta-analysis increases the sample size of an analysis. So, if Heneghan et al. were to defend the methodological decisions in the original study by saying that the underlying studies were value-apt, they would also have to concede that their own work was biased.

If Heneghan et al. (2019) are to defend the methodological decisions made in the underlying study, it must be, as it in fact was defended, on epistemological grounds (e.g., that thorough questioning is unlikely to create recall bias). However, to reiterate a crucial point, if this debate had occurred at the time the original studies were carried out and focused on a particular study and circumstance, the situation would have been different. In such a case, assuming one is willing to dispense with the value-free ideal, even if one knew with moral certainty that thorough questioning would introduce a nontrivial bias to the data, one might also have a very particular use in mind for the research and argue that thorough questioning is the best way to manage the inductive risk. In so doing, and assuming that the values are used judiciously and correspond to stakeholder values in the appropriate way, one is creating knowledge that is epistemically unimpeachable, given the relevant values, and is trustworthy for the intended users. What it would not be is knowledge that can be reliably used beyond the circumstances for which it was created. The knowledge is bespoke.

5. The trilemma of bespoke knowledge

In the previous section, I illustrated how a standard use of the argument from inductive risk had unforeseen consequences when value-laden knowledge traveled beyond the context for which the methodological choices had been tailored. This problem arose because of how the use of values affected the research produced. In this section, I wish to pull back from the specific details of the HPT case and state these issues more generally. In short, we have a set of three qualities that are typically seen as desirable but are often mutually incompatible: value-aptness, avoiding epistemic distortion, and universality (especially the ability of data to travel).

[6] For the purposes of this article, I will confine my discussion of bespoke knowledge to research that involves statistical analyses. Conceptually, bespoke knowledge is any knowledge claim that is tailored to a precise time and circumstance, and clearly, nothing in this definition is restricted to research involving statistical analyses. Nevertheless, because the issue emerges so clearly in statistical analyses, I have focused my attention here to illustrate the concept. I will leave it to future work to determine whether the same issue arises in nonstatistical contexts.

https://doi.org/10.1017/psa.2026.10232 Published online by Cambridge University Press

In some cases, one may be able to generate value-apt, reliable research that can travel by refraining from using values in ways that affect the measure of central tendency. But suppose the only way to create value-apt knowledge is by making methodological decisions that affect the measure of central tendency. If we adopt a lexigraphic ordering that prioritizes universality and epistemic reliability, we will have to settle for an inapt alignment between our science and our nonepistemic values. Such a position entails that, oftentimes, values cannot be allowed to influence methodological decisions. Although some philosophers endorse such "strict standards" (e.g., Steel 2017), most hold that values should have wider latitude to influence scientific decisions.

So, a second option is that researchers can produce value-apt research and allow it to travel. As a consequence, they must be resolved to introducing an epistemic distortion into the body of scientific knowledge.[7] This is arguably an unacceptable violation of scientists' epistemic obligations (Holman and Branch, unpublished). Indeed, in most cases, I suspect that neither of the first two options will be seen as desirable.

Accordingly, we have a final option. Scientists produce value-apt results that are epistemically reliable when confined. The consequence is that inquiry will produce knowledge that is tailored to specific circumstances and cannot travel. It will produce bespoke knowledge.

6. Conclusion

I began by noting that Holman and Wilholt (2022) identify three purposes that the value-free ideal was supposed to serve: veracity, authority, and universality. The former two have received significant philosophical attention (Wagner 2022; de Melo-Martín 2024). Indeed, de Melo-Martín has argued that all considerations relevant to proposing a solution to the new demarcation problem can be reduced to some combination of maintaining epistemic integrity and ensuring a politically legitimate selection of values. The argument made here shows that there are additional considerations.[8]

Again consider the HPT case study, but this time, let us take the vantage point of the original researchers (e.g., Nora and Nora 1975). Let us suppose we have good reason to believe that thorough questioning will introduce a bias into the data, and we are considering whether to use the method in spite of this concern. One consideration will be epistemological: How often will thorough questioning interact with HPT usage such that women who gave birth to a child with a congenital malformation are more

[7] A referee suggested that the introduction of bias can be mitigated by disclosure. My experience with industry-funded science leads me to a more skeptical position on the value of disclosure. If it prevents uptake, the research does not travel. If does not, either the bias is introduced, or we have some means of debiasing the data. I consider the burden of proof to be on the claimant that we can debias the data.

[8] In other work, T. Y. Branch and I argue that the desiderata are composed of at least three matched pairs of obligations and aspirations, where the obligations (epistemic distortion, value imposition, and undermining public trust) are imperatives that any replacement ideal must support, and the aspirations (veracity, universality, and authority) are objectives that are desirable but can be traded off against one another (Holman and Branch, unpublished). In our framework, a replacement ideal that countenances the production of bespoke knowledge weighs against an ideal, but it is a shortcoming that may be outweighed by other merits.

https://doi.org/10.1017/psa.2026.10232 Published online by Cambridge University Press

likely to remember that they took an HPT? That is, how big will the bias in the data be? A second consideration will be political: What are the values of the stakeholders in this research, and how can they be incorporated so as to produce trusted and trustworthy knowledge? That is, what do our democratic ideals dictate about scientists' role in knowledge production and public decision making? But beyond these epistemological and political considerations, there is a practical consideration: How important is it that our research is able to add to the public store of knowledge?

This final question is distinct from the previous two considerations. To wit, in section 4, we considered a case where all parties shared the same values and thus viewed the original studies as epistemically sound. Accordingly, the political issue does not arise because we consider a case where there is no contention about which values are the right values. We also stipulated that given those values, this was the correct methodology, so the epistemological issue is also satisfied. And yet, the practical issue of whether to generate data that could be used in the meta-analysis remains.

Even if Nora and Nora (1975) believed that the epistemological and political considerations justified the use of thorough questioning, the practical considerations could still have changed the decision. They may have chosen to conduct their research with the intent of providing definitive guidance for a defined set of stakeholders. Alternatively, the practical drawbacks of creating bespoke knowledge might lead them to reverse course. They might, for example, reason that their research will not be definitive, and to the extent that it is relevant to the regulatory decision of whether HPTs are removed from the market, it will only be as a part of the larger body of scientific knowledge.

References

Alexandrova, Anna. 2017. *A Philosophy for the Science of Well-Being*. Oxford University Press.

Anderson, Elizabeth. 2004. "Uses of Value Judgments in Science: A General Argument, with Lessons from a Case Study of Feminist Research on Divorce." *Hypatia* 19 (1):1–24.

Betz, Gregor. 2013. "In Defence of the Value Free Ideal." *European Journal for Philosophy of Science* 3 (2):207–20. https://doi.org/10.1007/s13194-012-0062-x.

Brown, Matthew J. 2013. "Values in Science beyond Underdetermination and Inductive Risk." *Philosophy of Science* 80 (5):829–39. https://doi.org/10.1086/673720.

de Melo-Martín, Inmaculada. 2024. "Concerns about Contextual Values in Science and the Legitimate/Illegitimate Distinction." *Philosophy of Science* 91 (4):851–68. https://doi.org/10.1017/psa.2024.20.

de Melo-Martín, Immaculada, and Kristin Intemann. 2016. "The Risk of Using Inductive Risk to Challenge the Value-Free Ideal." *Philosophy of Science* 83 (4):500–520. https://doi.org/10.1086/687259.

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

Elliott, Kevin C., and Ted Richards. 2017. "Exploring Inductive Risk: Future Questions." In *Exploring Inductive Risk*, edited by Kevin C. Elliott and Ted Richards, 261–78. Oxford: Oxford University Press.

Gal, Isabel, Brian Kirman, and Jan Stern. 1967. "Hormonal Pregnancy Tests and Congenital Malformation." *Nature* 216 (5110):83. https://doi.org/10.1038/216083a0.

Greenberg, Gillian, William Howard Wallace Inman, Josephine A. Weatherall, Abraham Manie Adelstein, and J. C. Haskey. 1977. "Maternal Drug Histories and Congenital Abnormalities." *British Medical Journal* 2 (6091):853–56. https://doi.org/10.1136/bmj.2.6091.853.

Heneghan, Carl, Jeffrey K. Aronson, Elizabeth Spencer, Bennett Holman, Kamal R. Mahtani, Rafael Perera, et al. 2019. "Oral Hormone Pregnancy Tests and the Risks of Congenital Malformations: A Systematic Review and Meta-Analysis." *F1000Research* 7:1725. https://doi.org/10.12688/f1000research.16758.2.

Holman, Bennett, and T. Y. Branch. Unpublished. "The Moral Terrain of the New Demarcation Problem: A Response to de Melo-Martín."

https://doi.org/10.1017/psa.2026.10232 Published online by Cambridge University Press

Holman, Bennett, and Torsten Wilholt. 2022. "The New Demarcation Problem." *Studies in History and Philosophy of Science* 91:211–20. https://doi.org/10.1016/j.shpsa.2021.11.011.

Irzik, Gürol, and Faik Kurtulmus. 2021. "Well-Ordered Science and Public Trust in Science." *Synthese* 198 (suppl. 19):4731–48. https://doi.org/10.1007/s11229-018-02022-7.

John, Stephen. 2019. "Science, Truth and Dictatorship: Wishful Thinking or Wishful Speaking?" *Studies in History and Philosophy of Science Part A* 78:64–72. https://doi.org/10.1016/j.shpsa.2018.12.003.

Nora, Audrey H., and James J. Nora. 1975. "A Syndrome of Multiple Congenital Anomalies Associated with Teratogenic Exposure." *Archives of Environmental Health: An International Journal* 30 (1):17–21. https://doi.org/10.1080/00039896.1975.10666626.

Olszynko-Gryn, Jesse, Eira Bjørvik, Merle Weßel, Solveig Jülich, and Cyrille Jean. 2018. "A Historical Argument for Regulatory Failure in the Case of Primodos and Other Hormone Pregnancy Tests." *Reproductive Biomedicine and Society Online* 6:34–44. https://doi.org/10.1016/j.rbms.2018.09.003.

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

UK Commission on Human Medicines. 2017. *Report of the Commission on Human Medicines' Expert Working Group on Hormone Pregnancy Tests*. London: UK Commission on Human Medicines.

UK Commission on Human Medicines. 2019. "Ad Hoc Expert Group on Systematic Review and Meta-Analysis of Studies on Oral Hormone Pregnancy Tests." https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/799874/CHM-Ad-Hoc-Expert-Group-paper_Heneghan-meta-analysis_March-2019_redacted.pdf.

Wagner, Wendy E. 2022. "No One Solution to the 'New Demarcation Problem'? A View from the Trenches." *Studies in History and Philosophy of Science* 92 (C):177–85. https://doi.org/10.1016/j.shpsa.2022.02.001.

Cite this article: Holman, Bennett. 2026. "Bespoke Knowledge." *Philosophy of Science*. https://doi.org/10.1017/psa.2026.10232

https://doi.org/10.1017/psa.2026.10232 Published online by Cambridge University Press