

Review of *Making Medical Knowledge*

Michelle Pham*†

Miriam Solomon, *Making Medical Knowledge*. Oxford: Oxford University Press (2015), xiii+261pp., \$60.00.

Introduction. Miriam Solomon's most recent monograph, *Making Medical Knowledge*, is about the social epistemology of the research and practice of Western medicine during the past 40 years. It provides a historical narrative of how medicine has been altered and negotiated by practitioners, with a focus on medical consensus conferences, evidence-based medicine, translational medicine, and narrative medicine. Solomon argues that some of the available literature on the epistemology of medicine tends to focus on just one method (e.g., narrative medicine), and the literature is not sufficiently critical. One aim then is to fill a gap in the epistemology of medicine literature about the variety of epistemic methods and how they interact to generate knowledge. This aim will provide a greater intellectual space to discuss normative questions about these methods and about the practice of medicine more broadly (10). Solomon argues that there is no single all-encompassing method for producing medical knowledge but instead a plurality of methods, which can come into conflict, or what she refers to as "untidy, methodological pluralism."

Overview. Solomon begins the book with the rejection of the "art" and "science" distinction in medicine. "Art" refers to the "soft" approaches to medicine, such as narrative medicine and work from the medical humanities, and "science" refers to the "hard" approaches, such as evidence-based medicine. Some proponents of the scientific approach unjustifiably rank those methods considered art to be lower than the scientific ones. This schema is also incomplete because there are methods that fall into neither category—for example,

Received December 2016.

*To contact the author, please write to: Department of Philosophy, University of Washington, Seattle, WA 98195; e-mail: mpham5@uw.edu.

†I thank Miriam Solomon for her comments and suggestions. I'm also grateful to Andrea Woody for her guidance and careful reads.

Philosophy of Science, 84 (April 2017) pp. 377–384. 0031-8248/2017/8402-0011\$10.00
Copyright 2017 by the Philosophy of Science Association. All rights reserved.

medical consensus conferences, which aim to achieve agreement by way of group deliberation. Solomon thus offers a different framework: “a developing, untidy, methodological pluralism” that does not oversimplify and hierarchically order the different methods of medical science. She focuses on the historical epistemology of medical consensus conferences, evidence-based medicine, translational medicine, and narrative medicine, which in turn support her thesis for untidy, methodological pluralism. I will briefly discuss each of these methods, except for translational medicine, and then explain how they figure in the argument for untidy, methodological pluralism.¹

Medical consensus conferences emerged in the 1970s in an effort to more quickly disseminate authoritative information between researchers and health care professionals. The US National Institutes of Health (NIH) model aimed to achieve “technical” (or “scientific”) consensus, that is, the resolution of scientific disagreements or controversies within the medical community through rational deliberation. Typically 10–20 medical experts participate, and the conference lasts 2–3 days.² The emphasis on group deliberation at these conferences prompted criticism, however, because expert and general consensus reached through such means may be caused by groupthink, peer pressure, or the anchoring phenomenon. Moreover, Solomon argues that technical consensus conferences merely endorse issues already agreed on by medical practitioners: “they typically took place *after* experts reached consensus, i.e., too late to produce a technical consensus” (55). These conferences, she argues, did not resolve scientific disagreements but were used to disseminate information. However, the NIH shut down their well-known program in 2013. These conferences still do operate, but they have dropped the use of “consensus conference” and now focus on evidence-based results rather than rational deliberation, given concerns about the influence of bias during deliberation.

Second, evidence-based medicine emerged in the early 1990s from clinical epidemiology. Solomon classifies this approach as a form of “*empiric*” medicine: the approach of figuring out which treatments work without necessarily understanding why. Randomized control trials are considered the highest form of evidence, with additional focus on other clinical trials, meta-analysis, and systematic review. This is contrasted with pathophysiological (or mech-

1. I will not discuss translational medicine since it is not central to Solomon’s case study on screening mammography (see chap. 7). Instead, one other method I will discuss is pathophysiological reasoning, which is not a new method from the past 40 years, but it is central to the case study on screening mammography. It also shares the same method with some translational medicine, i.e., T1 (155–59).

2. In comparison, the Danish model aims to achieve “interface consensus,” or critical discussion that includes the participation of laypersons and politicians. The goal is to have a diverse group of participants reach agreement so that they could act on a practical issue, such as crafting public policy.

anistic) reasoning, which is knowledge of how or why certain mechanisms work the ways they do. Proponents of evidence-based medicine do not place importance on the need to understand underlying mechanisms because, even with such knowledge, fruitful predictions for successful interventions often fail to materialize. To give some examples, “bone marrow transplantation for metastatic breast cancer, early attempts at gene therapy, angiogenesis inhibitors for cancer, and vertebroplasty for osteoporotic spinal fractures” were all reasonable approaches using mechanistic reasoning, but they did not deliver improved outcomes for patients (117). Despite the eschewing of scientific theorizing about mechanisms, Solomon points out that evidence-based practitioners still, nevertheless, use such information to develop interventions and to design randomized control trials. That said, pathophysiological knowledge is not necessary for evaluating whether new interventions are effective. Solomon, for example, cites the increased life expectancy for patients with cystic fibrosis through low-tech, evidence-based interventions such as chest percussion to loosen mucus, hypertonic saline mist, ibuprofen, and aerobic exercise (127). Knowledge about the genetic mechanisms of cystic fibrosis has indeed improved substantially, but it has yielded little success in terms of interventions, although some recent developments show promise (e.g., some cystic fibrosis mutations). Many of the successful low-tech interventions for cystic fibrosis have been inspired by looking at other conditions with similar symptoms.

A number of criticisms against evidence-based medicine are discussed in the book. I will briefly mention two. One is that evidence-based medicine’s emphasis on research interventions that work overlooks the unique and individual needs of patients—with respect to both their biological variability and personhood. What is further needed is the judgment of the clinician who bears a more intimate relationship with the patient and who in turn can tailor which interventions would work best for patients. This is also why clinicians have sometimes criticized evidence-based medicine as “cookbook medicine,” an approach that assumes interventions work in a “one size fits all” manner. Another criticism is that evidence-based medicine itself has not been subject to its own empirical standards, that is, systematic evidence reviews of whether research on interventions that work actually do improve overall patient care. Solomon points out that while in theory such interventions from evidence-based medicine should improve patient outcomes, the differences between places that practice evidence-based medicine and those that do not have not been empirically tested (151).

Third, narrative medicine developed recently from the medical humanities in response to “cookbook” (i.e., evidence-based) medicine. It emphasizes narrative “in the form of the patients’ story, the physicians’ story, or a story co-constructed by patient and physician” (178). The patients’ “psychological variability” or their unique “preferences and values” are considered valu-

able for physicians to make proper recommendations (178). Furthermore, the physician-patient relationship, rather than the physician acting as an authoritative expert who knows what is best for their patient, is crucial. The patients' needs are not determined purely by treatment of a biological problem but also with their voiced concerns through dialogue with their physician.

One conclusion to draw from the variety of methods discussed is that there is no unique method for generating medical knowledge but a plurality of them. Solomon points out that there are at least two ways to conceptualize this pluralism. The first, which Solomon calls "tidy pluralism," says that the methods should be conceptualized as a spectrum, organized as follows: pathophysiological reasoning (or translational medicine; phase 1 and 2 trials), evidence-based medicine (phase 3 trials, clinical trials, meta-analysis), medical consensus conferences (knowledge dissemination), and finally narrative medicine (knowledge implemented in clinical settings).³ In this picture there is a clear demarcation of where and when each method occurs, and they are each efficacious for the different kinds of questions they aim to answer (208). Each method contributes to the generation of medical knowledge in its separate domain. Solomon argues that tidy pluralism does not correctly reflect the pluralism we see in medicine, and she instead argues for a "developing, *untidy*, methodological pluralism" (my emphasis; henceforth, untidy pluralism). The methods in medicine are neither confined to particular stages of research nor only used to answer questions unique to their domains; rather, different methods can be used to answer the same questions and even provide conflicting answers, and there can also be conflict among practitioners working with one method.

Solomon illustrates medicine's untidy pluralism with a case study of the controversy about screening mammography for women between age 40 and 49. This case is assessed through recommendations from practitioners using the methods of (a) pathophysiological (causal, mechanistic) reasoning, (b) evidence-based medicine results, (c) expert consensus, and (d) clinical judgment (with narrative medicine). The causal evidence from pathophysiological reasoning is unclear since some believe screening mammography "causes breast cancer through the unforeseen effects of biopsy of suspicious lesions" (212). Evidence-based medicine results are unclear because there are harms as well as benefits of getting the procedure. On the consensus of experts, there is no uniform agreement from the different technical consensus conferences held on this issue—some recommend the procedure while others do not. Also noteworthy is that the participants at the NIH's

3. I focus on pathophysiological reasoning because it is central to Solomon's case study on screening mammography, although it shares a similar method with translational medicine. Both pathophysiological reasoning and translational medicine are classified by Solomon as belonging to the context of discovery (see 125 and 171).

consensus conference in 1997 could not reach agreement on this issue (212). And finally, clinician judgment in conjunction with the practice of narrative medicine is highly contingent on clinicians' expertise, as well as the patient's own narrative, although the procedure is generally viewed positively.

The efficacy of screening mammography has been controversial for over 40 years because there is no consilience among the different approaches mentioned. The recommendations vary, and there is further disagreement among experts working within each methodological domain. Solomon writes: "Despite the fact that mammography has been more extensively evaluated by randomized control trials than any other screening method, . . . it continues to be controversial, especially for routine use in women aged 40–49" (210–11). This case study, Solomon argues, shows that "[untidy,] methodological pluralism is [generally] responsible for much of the controversy over screening mammography" (217).⁴ Tidy pluralism, on the other hand, lacks the conceptual resources to account for a case in which the different methods are used to answer the same question but with conflicting results.

Untidy Pluralism. In the philosophy of science, pluralism (metaphysical or methodological) has had some notable proponents in the recent past, including Dupré (1993), Cartwright (1999), and Longino (2002). Solomon's book also advances support for pluralism but in the context of medicine. A central claim in *Making Medical Knowledge* is that untidy, methodological pluralism is a descriptively accurate theory about the methodological practices of medicine: they are varied and sometimes conflict when answering the same question(s). I take Solomon's descriptive account to provide foreground for thinking about normative questions that pertain to pluralism in medicine more generally. The attractiveness of this line of reasoning, however, depends on how well the proposed theoretical frame explains the practices of a scientific field. Solomon's case study on screening mammography shows how different methods can be used to assess the efficacy of a procedure, and they each provide "different answers and there is no 'meta-method' to which to appeal to resolve the disagreements" (208). In what remains, I will offer two points of discussion.

First, there may be room to challenge whether the domains of clinical judgment and narrative medicine can be used to answer the same question(s) as pathophysiological reasoning, evidence-based results, and consensus conferences. As Solomon details, clinicians who practice narrative medicine are interested in the medical conditions of each patient and his or her personal narrative. Clinicians' judgment encompasses not just biological and evidential content but also their interaction with particular patients, which makes their

4. Solomon cites a couple of other factors, which include political worries about decreasing funding for breast cancer research (217).

judgment about the efficacy of screening mammography for individual patients more nuanced and complicated. For example, the patient's values, social needs, and circumstances need to be taken into account when weighing the general benefit (or lack thereof) of screening mammography. In contrast, the question for the three other methods is a general research question applied to populations. This is not to say that there is no overlap between general research and clinician-patient contexts, quite the contrary, but this is much weaker than claiming that all of these domains can answer the same question.⁵ It is not clear to me, at least on the face of it, that grouping clinician judgment with the other three methods quite works since the information to be considered by clinicians includes much more than research results, and it is also because of this that we should not be surprised that there will rarely be a straightforward yes or no answer about the efficacy of screening mammography from clinicians who practice narrative medicine. This is a point that I think Solomon could clarify.

Second, there may also be room to explore whether technical consensus conferences are a method of the same sort as pathophysiological reasoning and evidence-based medicine. Technical consensus conferences are, in Solomon's words, "epistemic rituals" that disseminate information—they do not settle controversial scientific debates (12). Discussions about scientific content must be drawn from somewhere, and presumably it is from pathophysiological reasoning and evidence-based medicine (even if there is disagreement). Solomon, for instance, writes: "[technical consensus conferences] 'close the gap between research and practice' by disseminating evidence-based results in an authoritative manner" (83). So why are these conferences classified as an "epistemic method"? What aspects of technical consensus conferences make medical knowledge? I ask these questions because the trustworthiness of such conferences has declined over time: the NIH shut down its program, and practitioners of evidenced-based medicine consider the deliberative procedures used unreliable (111–12). The compressed procedural standards of technical consensus conferences have not, as Solomon points out, been especially viable, and this gives reason to question whether they should be classified as an epistemic method at all. What is it about consensus conferences that makes them categorically count as an "epistemological approach" or "method" (1)?

I wonder whether Solomon might respond as follows. Making medical knowledge is not simply about the evidential content, which might—naively—be considered the real epistemic content of medicine. If this were true, only pathophysiological reasoning and evidence-based results would be worth talking about. This view shares undertones with logical empiricism,

5. Note that one can make this claim without implicitly appealing to something like the art/science schema.

with its focus on the inferential relation between theory (or hypothesis) and evidence, which was rejected at the outset of the book (7, 11).⁶ On the contrary, making medical knowledge is an inherently social endeavor involving a plurality of approaches, and consensus conferences are as much a part of the process of medical knowledge production as are pathophysiological and evidence-based results. More specifically, technical consensus conferences' aim of disseminating knowledge is crucial to practice in a way similar to peer-reviewed journals, big databanks, and conferences, for they all promote information sharing among practitioners. The production of medical research (and more broadly scientific research) is a collective endeavor with many participants, and knowledge dissemination is required for it to work.

Surely scientific knowledge production—medical or otherwise—involves more than the inferential relation between theory and evidence. Even so, what gets classified as 'more' here does not come for free. The epistemic status of other features, crucial to practice, requires more than descriptive accuracy. Solomon's analysis shows that at best technical consensus conferences disseminate information in the vein of authority-based knowledge. And Solomon, at one point, notes that they should be evaluated "in terms of their broader epistemic achievements, especially the tasks of knowledge dissemination, practice change, maintenance of epistemic authority, and stimulus of particular kinds of research" (62). But why are these achievements considered a sort of epistemic approach or method? Specific to the case study, if the various technical consensus conferences are simply spreading information about the efficacy of (or problems with) screening mammography from pathophysiological reasoning and evidence-based results known to many practitioners, what is it about the procedural methods of consensus conferences that makes them epistemically distinct? I should stress that I take no issue with talking about the kind of role that consensus conferences generally have in the research and practice of medicine. Rather, I wonder why it is categorically an epistemic approach or method of the same sort as pathophysiological reasoning and evidence-based results. I suspect that Solomon may have more to say about this. But what Solomon has convinced me of most is that there is a plethora of problems with technical consensus conferences.

If my points are on target, then there is room to question Solomon's argument in favor of untidy pluralism. If clinician judgment and narrative medicine answer questions that are different in kind compared to pathophysiological reasoning, evidence-based medicine, and medical consensus conferences, then there may just be three epistemic methods that can be used to answer the same question. And if medical consensus conferences should not be classified as a method, then we may be left with two epistemic methods that can be used to answer the same question. Given these two points, one

6. See Richardson (2007) for a different reading of logical empiricism.

might doubt the viability of Solomon's untidy pluralism since there may really only be two methods that occasionally conflict—pathophysiological reasoning and evidence-based results—and not four.

Conclusion. My aim here was to provoke discussion about Solomon's untidy pluralism as a tenable theoretical frame to capture the epistemology of medicine, although there are many other interesting points of discussion to raise about the book. Historians, philosophers, practitioners, and sociologists of both medicine and science will find Solomon's *Making Medical Knowledge* to be an insightful read, especially her analysis of medical consensus conferences and evidence-based medicine. This is an interesting book, and I highly recommend it.

REFERENCES

- Cartwright, N. 1999. *The Dappled World*. Cambridge: Cambridge University Press.
- Dupré, J. 1993. *The Disorder of Things: Metaphysical Foundations for the Disunity of Science*. Cambridge, MA: Harvard University Press.
- Longino, H. 2002. *The Fate of Knowledge*. Princeton, NJ: Princeton University Press.
- Richardson, A. 2007. "'That Sort of Everyday Image of Positivism': Thomas Kuhn and the Decline of Logical Empiricist Philosophy of Science." In *The Cambridge Companion to Logical Empiricism*, ed. Alan Richardson and Thomas Uebel, 346–69. Cambridge: Cambridge University Press.