Laszlo Kosolosky
Centre for Logic and Philosophy of Science
Ghent University

The intended window of epistemic opportunity: a comment on Miriam Solomon

Abstract: In this paper, I argue that Miriam Solomon fails to show that medical consensus conferences, as organised by the National Institute of Health (NIH), miss the intended window of epistemic opportunity (Solomon, 2007: 170), and accordingly take place after the experts have reached consensus. This is done, on the one hand, by differentiating between, what I call, 'academic' and 'interface consensus', and, on the other hand, by analyzing the arguments and argumentation style Solomon uses to make her claim explicit. At the very least, the overall argument suggests that her statement is inadequately supported, if not that the opposite claim is true. In this manner, I intend to provide additional insight into the notion of 'consensus' when applied to scientific practice. Any decisions as to change the NIH consensus development program should take these actual achievements into account.

1. INTRODUCTION

In a nutshell, Miriam Solomon defended in her book chapter 'The social epistemology of NIH consensus conferences' (2007) that the National Institute of Health consensus development conferences do not bring about rational consensus on controversial health topics. Solomon insists that a consensus usually exists beforehand, at least among the researchers, as opposed to what the NIH claims to be the case. According to her, these conferences can only serve a (subsidiary) goal of spreading the information across to a more general public.

In the first part of the paper, I look into the NIH consensus development program. In the second part, I differentiate between 'academic' and 'interface consensus', which enables me to show where our views differ. The third part analyses her reasoning by focussing on (1) the argumentation style and (2) the arguments presented. Together these parts point out that Solomon fails to show that *"consensus conferences miss the intended window of epistemic opportunity"* (Solomon, 2007: 170). She fails to indicate that consensus conferences typically take place after the experts reach consensus.

2. NIH CONSENSUS DEVELOPMENT PROGRAM

The NIH consensus development program constructs major conferences that produce evidence-based consensus statements addressing controversial issues in medicine important to health care providers, patients, and the general public. The conferences aspire to provide an independent look at the issues through an unbiased panel. These conferences are run on a type of 'court model', in which the panel members are like a jury. They have no financial or career interests related to the topic and they are highly regarded in their own fields, but are in no way closely aligned with the subject.

The review-process itself is structured as follows: First, there is an in-depth presentation of evidence to the panel. This includes a systematic literature review[1] prepared by the Agency

[1] A systematic review is a literature review that tries to identify, appraise, select, and synthesize all high quality research evidence relevant to a particular research question.

for Healthcare Research and Quality (AHRQ). In addition, recognized experts on the topic give presentations to the panel and the audience. Finally, formal periods of public discussion are held. The conference program contains approximately 21 speakers: 3 of them present the information found in the systematic review of the literature; the other 18 are experts on the topic at hand, have likely published on the matter, and may have strong opinions or beliefs on the topic. Crucial here is that where multiple viewpoints on a topic exist, every effort is made to include speakers who address all sides of the issue.

Due to the stress on providing an evidence-based consensus, the NIH differentiates between Consensus Development Conferences and State-of-the-Science Conferences. Consensus Development Conferences are typically undertaken when there is a solid body of high-quality evidence, such as randomized trials and well-designed observational studies. State-of-the-Science Conferences are generally utilized in cases where the evidence base is weaker. In both cases the statement is a report evaluating scientific information on a given biomedical or public health intervention with the purpose of resolving a particular controversial issue in clinical practice. Each report handles a series of five to six questions concerning efficacy, risk, clinical applications, and directions for future research.

3. ACADEMIC VS. INTERFACE CONSENSUS

As Solomon points out, and as sketched out in the NIH guidelines, these conferences have two goals: On the one hand, they aspire to bring about rational consensus on controversial health topics, and, on the other hand, they intend to spread the medical information across to the broader public. Now, according to Solomon, these conferences do not meet their first goal, because of consensus already existing. They only serve a (subsidiary) goal of spreading the information. In her words, consensus conferences can be merely a *"[…] rhetorically efficacious way to get the word out, to interested intermediaries such as professional groups, pharmaceutical companies and health insurance companies who will then adapt the statements for their own particular purposes"* (Solomon, 2007: 175).

Before looking further into her position, I introduce an analytic distinction in dealing with consensus making[2], which helps to point out where exactly my position differs from Solomon's. Talking about consensus in general implies supporting a distinction between consensus-making among scientists – establishing an academic consensus – on the one hand and consensus-making at the interface between science and society – establishing an interface consensus – on the other hand. The NIH conferences carry the task of contributing both to 'academic consensus' as to 'interface consensus'. These conferences contribute to academic consensus by establishing a consensus within the scientific community and to interface consensus by providing and transmitting this established scientific consensus to the larger community. Underneath are two moments of decision-making that should be analytically separated: the actual move from plurality to consensus among scientists or experts[3] and the moment of dissemination and justification towards society[4]. To get an (intuitive) grasp of the distinction at hand it suffices (1) to look into the particular relationships at play between the actors in each type of consensus, and (2) to assess the conditions that need to be fulfilled beforehand.

[2] This distinction can be supported irrespective of the view one has on consensus itself, i.e. whether one sees it as the instance where dissent amounts to zero (Solomon's position) or where everyone unanimously agrees or what have you. For further insights on the notion of consensus, see Kosolosky & Van Bouwel (in preparation).
[3] Which according to Solomon happens outside of consensus conferences.
[4] Obviously, there is overlap between both moments and there might be causal influences in both directions.

The relationship at play within academic consensus is one between experts. In the academic world, every scientist/academic is regarded to be an (equal) peer and everyone serves as an authority within his or her field. These people are generally regarded to be on the cutting edge of research and are expected to be among the first to notice changes occurring within their field of expertise. The relationship at play within interface consensus is one between expert and layman, grasping the interface between science and society. The NIH consensus reports contribute to an interface consensus by providing an assessable source of information for the public, namely the consensus reports. This type entails a relation between expert and layman grounded on authority, trust, and mutual respect, where the actors are not regarded to be on equal footing. The difference in interaction is important to bear in mind when we want to have a look at what's at stake in each of them. This brings us to our second point, namely the conditions that need to be fulfilled beforehand. As for academic consensus it was already pointed out that enough evidence has to be available on the basis of which a conclusion can be formulated. As for interface consensus, an academic consensus is needed if one wants NIH consensus conferences to be able to bridge the gap between theory and practice. Transmitting this academic/scientific consensus in an understandable manner to the public will allow one to act accordingly. So a necessary condition for being able to contribute to interface consensus is there being an academic consensus[5].

Although Solomon will agree that a consensus conference can contribute to interface consensus, she will not agree that consensus conferences contribute to academic consensus, because of it already existing beforehand. According to her, consensus conferences miss the intended window of epistemic opportunity: they typically take place after the experts have reached consensus. In the following section of the paper, I investigate whether the arguments she does present to ground her claim upon are conclusive or not.

4. REFUTING SOLOMON'S ARGUMENTS

Particularly baffling is the way in which Solomon argues for such a quite substantial claim. If one gazes upon the history of NIH consensus conferences one can see that the NIH already organised (approximately) 157 conferences. Solomon discusses two of them in further detail, which are supposed to serve as central examples to clarify what NIH consensus conferences actually entail. However, when push comes to shove, she omits to say why these two are typical examples. So as a more general remark, the question remains whether they really are as typical as she supposes them to be and on what grounds she (can) make(s) this assumption. If she aspires to argue that (academic) consensus usually exists beforehand, she will have to make further (and more substantial) efforts.

Solomon defends her view by presenting three arguments. In this section, I address each of them and see whether they hold up to scrutiny:

1) According to Solomon, the 'Helicobacter Pylori in Peptic Ulcer Disease conference' (1994) *"took place after the important clinical trials […] and after research scientists, and many*

[5] This is at least the case when talking about consensus development conferences, as Solomon seems to be doing. As for state-of-the-science conferences, contributing to academic consensus is done by sketching out on which topics there is agreement and on which there is none (and thus further research is required). Contributing to interface consensus is less obvious here, unless formulated in a negative manner by stating that there is not enough evidence (and agreement) present to make positive recommendations towards the public in dealing with a certain disease.

prominent clinicians, had reached consensus on the use of antibiotics for peptic ulcers" (Solomon, 2007: 170). The only source she mentions for this claim is the report issued by the NIH[6]. This report reflects all the scientific studies that established a disturbing epidemiologic relationship between H. pylori and gastric malignancies, concluding that *"such studies have given rise to the hypothesis that H. pylori is a major etiologic factor in peptic ulcer disease and that diagnosis and eradication of the organism are necessary for optimal therapy of the disorder"* (NIH, 1994: 3-4). The report says nothing on there being a consensus already established. It only mentions there being sufficient evidence to look into the matter more thoroughly (*hypothesis*), which is, as sketched out, a crucial requirement for these evidence-based consensus conferences to be organised. Solomon's claim of there already being an established consensus is unsound, unless she wishes to equate consensus with 'scientific studies being published that together seem to point in a particular direction', which is clearly insufficient. The fact that a couple of peer reviewers agree that a scientific study should be published does not entail there being an academic/scientific consensus[7]. Even if these studies do not raise criticism from experts who actually read the studies, one can surely imagine other experts not having read the studies and, thus, not having been informed, which a priori excludes there being consensus amongst academics beforehand. This is mainly because it has not been determined whether these experts would agree with the published results or not. There is even a possibility (1) that researchers claim exactly the opposite based on the studies they have read or undertaken themselves, or (2) that they do not find the available evidence convincing enough.

2) According to Solomon *"the 2002 Consensus Development Conference 'Management of Hepatitis C: 2002' repeats recommendations that were already stated by the FDA in the previous year"* (Solomon, 2007: 170). The two 2001 studies dealing with hepatitis C issued by the Food and Drug Administration are: 'FDA "Ribavirin and chronic hepatitis C infection", Consumer, 2001; 35(5): 3' and 'Schwetz B.A. From the FDA, JAMA, 2001; 286(10): 1166'. What both studies acknowledge is the fact that the FDA has issued two approvals involving the use of Rebetol capsules (ribavirin) to treat patients with chronic hepatitis C. Now if we look at the outline of the final NIH report, we notice that the report deals with the following five questions: (1) What is the natural history of hepatitis C?, (2) What is the most effective appropriate approach to diagnose and monitor patients?, (3) What is the most effective therapy for hepatitis C?, (4) Which patients with hepatitis C should be treated?, and (5) What recommendations can be made to patients to prevent transmission of hepatitis C? (NIH, 2002: 7). In response to the question that comes closest to the one being answered by the FDA report, i.e. 'What is the most effective therapy for hepatitis C?', the NIH report says that *"combination therapy results in better treatment responses than monotherapy, but the highest response rates have been achieved with pegylated interferon in combination with ribavirin. […] Currently the best indicator of effective treatment is an SVR, […]"* (NIH, 2002: 17). A crucial nuance is at stake here: whereas the FDA reports talk about an appropriate method of dealing with hepatitis C, namely taking ribavirin, the NIH report addresses the question what the best (or most effective) therapy for hepatitis C is. An FDA report does not address the question of most effective therapy, it merely *"[…] requires drugs to be tested only relative to placebos. This means that an FDA approval is, at best, a signal that the approved drug is better than taking a sugar pill, not that it's better than an existing treatment"* (Reiss, 2010: 9). Moreover, the other types of questions the NIH report dealt with were not addressed in any of the FDA reports. Taken all together, the NIH report displays

[6] A systematic review undertaken by the AHRQ did not exist at that time (only from 2001 onwards). So this type of evidence cannot be used to justify the claim made here.

[7] For further insights on the ambiguity of peer review, see Kosolosky (forthcoming).

something more substantial than merely repeating FDA recommendations, it takes these recommendations to another level and incorporates them in a larger framework. So again Solomon did not succeed in showing that consensus already existed beforehand.

3) As for her final argument, she refers to a quote by John Ferguson, a long-time director of the consensus development program: *"often the planners of any given consensus conference are aware of a likely outcome and use the conference as a mechanism to inform the health care community"* (Ferguson, 1993). The statement is insufficient evidence for her claim on consensus, for at least two reasons: (1) a likely outcome is not necessarily the actual outcome, and (2) it does not show that there already existed a consensus within the scientific community. Ferguson's quote only expresses a personal opinion on the matter based on own experience. If Solomon wants to claim that consensus usually exists beforehand, she needs to present systematical research data of past conferences that supports her conclusion. Unless John Ferguson has in fact undertaken this research project or has other substantial evidence to support his claim with, one can only ascribe this argument the credit it deserves, namely an authority argument. As long as Solomon does not clarify what type of evidence John Ferguson used to reach this conclusion, we cannot regard this statement to be more than that, a statement.

My analysis shows that Solomon's arguments are inconclusive, but this in no way entails that the position she defends - consensus preceding consensus conferences - is false. One could argue that it is still the case that scientific consensus exists beforehand, but that Solomon just failed to argue for it in an appropriate manner. What it does show is that some 'central' examples of consensus conferences turn out to be examples arguing for the opposite claim, i.e. that scientific consensus does not precede consensus conferences. Moreover it shows that a clear notion of consensus that corresponds with the one the NIH has in mind is needed if one wants to make a substantial claim on the matter. The NIH consensus development program and the dual goals set out by the NIH stress the necessity of a stable consensus. As for the NIH, this appears to entail that deliberation and critically sharing of available information are two conditions that play a decisive factor in consensus making. Any account of when scientific consensus came to be, without a clear notion of scientific consensus itself, is thus doomed to fall short[8].

5. CONCLUSION

In this brief reaction to Miriam Solomon, I have been able to show that the central claim she makes in her book chapter is at the very least inadequately grounded. Through an analysis of her argumentation style and corresponding arguments, it has become evident that she did not succeed in her design. In general, she did not show what 'scientific consensus' entails in the NIH Consensus Development Program. In particular, that scientific consensus precedes consensus conferences is clearly not shown to be the case. I opted for a more thorough way of dealing with consensus in scientific practice, by making a difference between 'academic' and 'interface consensus', which allowed me to analyse the full extent of these consensus conferences and the interaction between the two main goals. Research on the notion of consensus itself and how it is used (and perceived) by the NIH is needed, if one wants to make convincing claims on the matter. The NIH seems to aspire a stable consensus and the

[8] Because I do not make the opposite claim here that consensus precedes consensus conferences, I leave out a full definition of consensus. An onset was given though, as well as the ground that needs to be covered if one wants to make a claim at all. For further insights on the benefits of consensus conferences and on (the stableness of) scientific consensus, see Kosolosky (under review) and Kosolosky & Van Bouwel (in preparation).

question remains whether the notion of consensus they use corresponds with the one Solomon has in mind.

Acknowledgements. I would like to thank Mathieu Beirlaen, Mark Brown, Heather Douglas, Jan De Winter, Kevin Elliott, Raoul Gervais, Boaz Miller, Miriam Solomon, Jeroen Van Bouwel and Erik Weber for their comments on earlier drafts of this paper. The research for this paper was supported by the FWO Flanders through research project nr. G.0122.10.

References

- Agency for Healthcare Research and Quality. [On-line]. Available: http://www.ahrq.gov
- FDA (2001). Ribavirin and chronic hepatitis C infection, Consumer, 35(5): 3.
- Ferguson J. (1993). NIH consensus conferences: dissemination and impact. Annual New York Academy of the Sciences, 703, 180-198.
- Kosolosky L. (under review). Consensus formation: introducing a continuum ranging from consensus conferences to meta-analysis in dealing with consensus driven aspirations.
- Kosolosky L. (forthcoming). "Peer review is melting our glaciers": Exploring how and why the Intergovernmental Panel on Climate Change (IPCC) went astray. Special issue of the Journal for General Philosophy of Science.
- Kosolosky L. & Van Bouwel J. (in preparation). Explicating ways of consensus making in science and society: a social account of consensus formation.
- National Institute of Health. Website on the Consensus Development Conference Program. [On-Line]. Available: http://www.consensus.nih.gov
- National Institute of Health. (1994). Helicobacter pylori in peptic ulcer disease. NIH Cons Statement. February 7-9; 12(1): 1-22. [On-Line]. Available: http://consensus.nih.gov/1994/1994HelicobacterPyloriUlcer094PDF.pdf
- National Institute of Health. (2002). Management of hepatitis C: 2002. NIH Consensus Statements. June 10-12; 19(3): 1-46. [On-Line]. Available: http://consensus.nih.gov/2002/2002HepatitisC2002116PDF.pdf
- Reiss J. (2010). In favour of a Millian proposal to reform biomedical research. Synthese, Online first, September 8, 2010.
- Schwetz B.A. (2001). From the FDA, JAMA, 286(10): 1166.
- Solomon M. (2007). The social epistemology of NIH consensus conferences. In H. Kincaid & S. McKitrick (eds.), Establishing medical reality: essays in the metaphysics and epistemology of biomedical science, Springerlink, 167-177.