Explicating different ways of consensus making at the interface between science and society

In our society, there are these moments in which establishing a *scientific consensus* seems imperative to solve urgent problems, for instance, as concerns climate change; achieving consensus on the causes and extent of global warming would facilitate policymaking and, moreover, send a convincing signal that doing nothing will have dire consequences. On the other hand, philosophers studying plurality and heterodoxy in science have raised questions concerning the ideal of the scientific consensus and the pernicious effects the consecration of scientific consensus might have.

Several philosophers of science have developed interesting accounts on how to deal with this tension between plurality and consensus; how scientific plurality, dissent, and consensus-making can go hand in hand: and, in relation hereto, how consensus conferences manage to deal with the tension between plurality and consensus. In what follows, we want to analyse some of these accounts about how science goes or should go from scientific plurality to scientific consensus, introducing an important 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.

This analytical distinction will help us (a) compare the differences among philosophers in understanding consensus-making, (b) develop our social-procedural account of scientific consensus (analogous with social accounts of objectivity) paying attention to *meta-consensus*, and (c) spell out how to modify the public understanding of science and the public use of the ideal of consensus.

1. Aspects of *scientific consensus*: the intra-scientific consensus and the science-society interface.

If we aim to elaborate on the tension between consensus and plurality mentioned above, we need first establish what aspects of scientific consensus we would like to address here. We draw a distinction between a 'technical, academic consensus' and an 'interface consensus'[1]. The former points at a consensus being established among scientists or experts in a certain field related to a certain topic. The latter relates to a consensus being established at the border between science and society, typically including a wider range of actors apart from scientists (i.e. laypeople, interactional experts, government representatives, etc.). The most intuitive way to grapple this distinction is by distinguishing two problems or two moments of decision-making that should be analytically separated, i.e. the actual move from plurality to consensus among scientists or experts on the one hand and the moment of dissemination and justification towards or within society on the other hand.

[1] We grant at once that talking about consensus can at the same time apply to academic as well as interface consensus. The distinction is meant to be non-exclusive. What we seek to elaborate are ideal types of consensus that can be used to classify or evaluate the plethora of non-ideal or mixed cases that dominate the research practice. These moments might de facto coincide, but might be distinguishable, e.g., as concerns the way uncertainties are weighted.

Another way to comprehend the distinction is by describing it in terms of the relation between its actors and/or the conditions that need to be fulfilled beforehand. 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 the interface consensus is one between expert and layman, grasping the interface between science and society. 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, we could argue that the community should enable critical interaction among academic experts and have significant evidence available on the basis of which a conclusion can be formulated. As for interface consensus, some form of academic consensus or a first attempt to establish consensus is required to start up the interface process.

A helpful example to illustrate how this distinction might be used are consensus conferences. To demonstrate this, we briefly sketch the workings and goals of the National Institute of Health Consensus Conferences (NIH): The NIH consensus development program constructs major conferences that aim to 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 serve as a jury. They are supposed to 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 under scrutiny.

Due to the stress on providing an evidence-based consensus, the NIH itself 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. As sketched out in the NIH guidelines, these conferences have set two goals for themselves: 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. We could say that the NIH conferences carry the task of contributing both to 'academic consensus' as to 'interface consensus', to the former by establishing a consensus within the scientific community and to the latter by providing and transmitting this established scientific consensus to the larger community (NIH, website).

Obviously, as thinking the example through shows, there might be an overlap between both moments and there might be causal influences in both directions. In the following sections, we will show how our analytical distinction helps in explicating different understandings of the structure and functions of consensus-making in the philosophical literature. First, in section 2, it will made clear how the example of the NIH consensus conferences already captures differences among philosophers in understanding consensus making.

2. Consensus-making conferences: Miriam Solomon on the two moments of consensus-making.

Miriam Solomon (2007) argues 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.

Put bluntly, 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 the latter consensus already existed before the conference started. According to her, consensus conferences miss the intended window of epistemic opportunity: they typically take place after the experts have reached consensus.[2]

In her understanding of the role of consensus conferences, Solomon (implicitly) distinguishes between the two moments identified above. She claims that the actual (NIH) consensus conferences should be seen as a moment of disseminating knowledge, justifying a clinical practice, … not as a place for doing science and establishing an intra-scientific consensus, that already happened beforehand. The consensus conference is then a ritual, a choreographic epistemic performance, creating authoritative knowledge; it is the moment of the *interface consensus* addressing the public intended. In her own 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).

Her account, however, is left with a couple of open problems. First, the way in which Solomon argues for such a quite substantial claim raises questions. If one gazes upon the history of NIH consensus conferences one can see that the NIH already organized over 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 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. Moreover she does not address how exactly the establishment of the intra-scientific, academic consensus came to be; Does establishing an

[2] In previous work, we analyzed her claim and found it to be wanting, mainly because the examples she gives in favor of her account do not suffice (cf. [AUTHOR], 2012).

academic consensus go further than published studies pointing in the same direction? And if the consensus conference is merely a ritual, if the only epistemic surplus/value is that it authorizes knowledge, then why was there a different outcome in NIH consensus conference 1977 and the Danish consensus conference 1983 (both on breast cancer screening)?[3] Do we have the same academic consensus (preceding consensus conferences) in both cases and a different interface consensus (as a result of the consensus conferences) because of the difference in publics intended?[4]

We will return to this matter below. In this section, we presented the role scientific consensus plays in consensus conferences and the way in which consensus is established, whereby two moments of consensus-making can be distinguished. According to Solomon, consensus conferences serve to the second one, i.e. the elaboration of an interface consensus, while the first one, the making of the academic consensus, happens beforehand (cf. Solomon 2007: 173, "the rational basis for that consensus is made clear by AHRQ formal assessment of that evidence" (AHRQ is the *Agency for Healthcare Research and Quality*).

3. Experts and consensus-making: John Beatty on the two moments.

Whereas Solomon's work teaches us a lot on the "interface moment", John Beatty's focus is in the first place on the scientific experts establishing a consensus among themselves, the *academic consensus*. Obviously, the possible societal impact and consequences are taken into account – leading to some kind of self-censorship (cf. Beatty, 2006) – but there is no consideration of an explicit *public* moment, during which an *interface consensus* would be established.

In his earlier work on experts (2006), Beatty already captures part of the tension between consensus making on the one hand and intrinsic value of plurality on the other hand. Through a case study of the maximal admittable dose of radiation for humans, he utters that craving consensus might widen the gap between expert and novice in two distinctive ways. First, what he calls simplification, entails that scientists instead of simply telling us what they know, they might tell us simply. In this manner, a lot of crucial information gets lost along the way. Second, what he identifies as the intentional withholding of information, means that scientists often agree amongst one another as to retain information from the public or silence discussions, which results in a distorted view of consensus amongst the public and nourishes the ill-conceived expectations they might have. Both pitfalls find their reasoning on either paternalistic or protective grounds, as Beatty elaborates. The former meaning that experts state that it might actually be in the public's advantage if they speak with one voice

[3] The first Danish medical consensus conference was held in 1983, on the topic of early detection of breast cancer. It did not recommend general mammographic screening, even for women over 50 years of age, thus coming to a different conclusion from the first NIH Consensus conference held in 1977 (see Jorgensen 1990). "The role that consensus conferences can play in making policy explains why different countries hold consensus conferences on the same topics. For example, recommendations for mammograms to screen for breast cancer may be different in countries with different resources or different values. For instance, in the UK, mammograms are not routinely offered to elderly women because they are not cost effective at that age; in the USA cost effectiveness is not valued so highly." (Solomon, 2011: 248)

[4]

rather than with many, whereas the latter hints at the fact that experts, in this manner, could in fact guarantee that their status remains intact and can prohibit others from gaining the authority and trust to do their work.

Building on these striking insights on how scientific consensus is actually achieved, Beatty (together with Alfred Moore) works out an account of consensus formation that is partly procedural and partly substantive[5], i.e. a notion of *deliberative acceptance*. In their paper 'Should we aim for consensus?' (2010), they start from an account of consensus-making set forth by Margaret Gilbert (1987). Gilbert captures a notion, of what she calls, joint acceptance: "A group jointly accepts *p* if and only if the individual members have openly agreed to let *p* stand as the position of the group" (Gilbert, 1987: 194). Beatty and Moore emphasize two characteristics of this account of consensus. First, consensus understood in this sense tackles the notion on a rather different level: there is no talk of consensus on a certain proposition *p*, but consensus to let *p* stand as the position of the group. On *p* itself there could be considerable disagreement. Let us clarify this with a simple example. Suppose your faculty board wants to hire a new full-time professor in ethics, and after discussion and voting they opt to go with candidate A. You, however, preferred candidate B, based on the previous experiences she acquired and the impression she gave during the interviews. Candidate A however is put forth as the department's choice. Now, in the same manner you can disagree as an expert on *p*, but yet again acknowledge the consensus recommending *p* (Beatty & Moore, 2010: 206). Moreover, in Gilbert's terms, one is thought not to publicly disagree with the group opinion, of course one is not expected to lie, but any dissenting views should be rather carefully expressed (Beatty & Moore, 2010: 207). Second, according to Beatty and Moore, defining consensus as such renders the notion meaningless. One can endorse a rather weakly defined notion as the consensus opinion, on which everyone agrees, but in such a case it is nothing more than an empty shell (Beatty & Moore, 2010: 208).[6]

Beatty and Moore, develop their own notion, partly inspired by Gilbert, of consensus formation, i.e. *deliberative acceptance*: "A group deliberatively accepts *p* if and only if the individual members, based on the quality of their deliberation, have openly agreed to let *p*

[5] In section 4 we clarify in detail what they mean by both parts of the definition.

[6] An historical example arguing in favor of this remark is the consensus report (and table) published by Peter Mitchell to support his chemiosmosis hypothesis over the dominant chemical or direct-interaction coupling theory, to explain how and why ions move from an area of high concentration to an area of low concentration. Mitchell, however, presented a distorted view of consensus in the scientific discipline in 3 respects: (1) He does not identify all the expert members dealing with this topic, as this better fits his view of consensus on his hypothesis growing steadily over the years. (2) He ascribed some of the experts a wrong position. For example, the table mentioned experts being already halfly convinced that his route was the way to go, whereas it turned out they were actually convinced of his approach being half right. (3) He omits to specify the full extent of the cognitive content of consensus and forgets to show how it coincides with the views of the scientists. The consensus view he put forward seemed to be fairly evident, whereas upon careful examination it turned out that every scientist fills in the definition of chemiosmosis rather different. The consensus approach was thus based on a basic version of chemiosmosis, however, "the scientific content of this basic version was minimal; so much so that it could be and frequently was endorsed by those who described themselves and were described by others as strongly opposed to chemiosmosis as well as by those who claimed to be fervent supporters of the theory" (Gilbert & Mulkay, 1984: 130). With Gilbert's understanding of consensus as joint acceptance, you might obtain situations like the Abilene paradox, where no one wants to "rock the boat", and, eventually, the consensus position that is accepted is counter to the position of every individual.

stand as the position of the group" (Beatty & Moore, 2010: 209). As opposed to joint acceptance, this view allows for an explicit moment where dissenting opinions could be heard and noticed through deliberation (expanded with a system of voting when necessary). Beatty and Moore are convinced that their approach surpasses the difficulties set forth for Gilbert's account: "Deemphasizing consensus on substantive issues (though not entirely), and stressing consensus on deliberative quality would not only take away the temptation to hide a persistent minority position, but would instead provide a good use for it. What better way to inspire confidence in a deliberative outcome than to show that 1) the position in question had been tested against a worthy alternative; 2) the minority felt that they had been heard, that they had been treated as deliberative equals; and 3) having been heard, even the minority agreed to let the position in question stand as the group's" (Beatty & Moore, 2010: 209). However, we think their account faces a couple of open problems. First, 'let the position stand' is too passive to be able to capture the interaction between opinions and dissenting views. This phrasing omits the fact that in scientific and interface discussions there are not only persistent minority positions (sometimes there even might not be any position that convinces more than half of the participants), but in most occasions continually present minorities who are not that eager to just give in. This relates to our second objection, i.e. the position of the minority. Deliberative acceptance in a sense is aiming too high. According to this view, there will be a time when the minority has to admit that they were heard and that some kind of epistemic justice has occurred, that they were found to be unconvincing, and that they should correspondingly endorse the consensus position of the group. Although Beatty and Moore spell out this objection themselves, they find it unconvincing to aim lower (Beatty & Moore, 2010: 209). Finally, moving from an academic consensus to an interface consensus, opens up the question what the responsibilities of the scientific expert are and what the responsibilities of society or policy-makers are, i.e. those that set-up expert committees and decide upon their functioning. As sketched out above, Beatty and Moore's account rather functions on the level of academic consensus. Interesting to see, however, would be how their account tackles the problems raised above.

Summarizing, Beatty & Moore focus on the formation of the academic consensus, keeping an eye on how the consensus could be sold to the public by the experts, and taking into account the expectations of the public (cf. section 6), but there is no consideration of an explicit *public* moment, during which an *interface consensus* would be established. Neither is the epistemic contribution of interactions between the academic and the interface moment explored (an interaction in which minority positions might play an important role, cf. below).

4. Consensus-making and the meta-consensus I: a social-procedural account of scientific consensus.

In developing their account of consensus, Beatty & Moore rely on democratic theory, just like we have done before (see [AUTHOR], 2009). We think this points at an important characteristic of consensus-making, in solving the tension between plurality and consensus, which is not always made explicit in accounts of consensus as knowledge-based: in both moments, there is a meta-consensus or a meta-agreement in play. Therefore, instead of focusing on consensus on the simple level, that is, as the result of alternative

theories/models tested against one another eventually – thought to be – leading to some consensus *outcome*, we could learn a lot by shifting to the analysis of the meta-consensus that stipulates the *procedure* to be followed. Understandings of consensus-making differ in how much weight they place on procedures relative to substantive considerations about the quality or characteristics of the outcomes of these processes. Let us elaborate this distinction further:

(a) Focusing on *outcome*, the substantive approach to consensus-making: Researchers agree on what they agree (in its ideal form mostly understood as a unanimity in which every participant comes to hold the same position for the same reasons) and the outcome satisfies certain conditions to assure that the consensus is right/correct and/or knowledge-based. Thus, there is a defined desirable outcome or decision (the outcome has to be right, living up to procedure-independent qualities), and in that sense, an implicit presumption of what a collective process or procedure has to generate qua collective decision; the existence of a consensus is presupposed. In this account, complete unanimity is not required, but only that if dissent exists, it is marginalized and suppressed. (Miller, 2013)

(b) Focusing on the *procedure*, the procedural approach to consensus: Researchers might agree on the procedure through which to aim for a collective decision/consensus. Call it a form of meta-agreement or meta-consensus, without specifying requirements on the substance, on the outcome, that this procedure eventually should produce. Here, there are no procedure-independent criteria to assess the quality of the correct outcome; it is the procedure itself that matters and that has to satisfy certain normative criteria. An outcome is legitimate if the collective decision-making procedure itself satisfies certain conditions and these conditions (the appropriateness of the procedures) might be satisfied in degrees (cf. infra) (Longino, 2002, chapter 6).

(c) Focusing partly on *procedure* and partly on *outcome*: Both the procedures and the outcome should be correct. Beatty & Moore's account presented in section 3 is a nice example: There is reference to substance in the act of accepting or letting the outcome stand. This acceptance might than be "based on the quality of your deliberation", which seems to bring it very close to a procedural view, but nevertheless omits to go all the way. According to this view, you might associate the idea of a substantive consensus with the idea of normative unanimity, that is, everybody coming to hold the same position for the same reasons. In much deliberative theory (Mansbridge, 2010: 66 + 68), it is thought that if a deliberative group cannot reach consensus in this sense, then the proper role of deliberation is simply to clarify and structure the disagreement, which sets the scene for a decision by some other (non-deliberative) method, usually a vote. Now what Beatty & Moore outline in the paper with the idea of joint acceptance and deliberative acceptance is a kind of consensus that is less than normative unanimity (in which scientists can speak as one without having to say that they were in full agreement), but more than simply the preferences of the participants combined under some decision rule, which need have no reference at all to substance.[7]

[7] We thank Alfred Moore for clearing this out.

Thus, the solution for the tension between plurality and consensus could be sought in some form of meta-consensus or meta-agreement. The resulting account of consensus will be a social one – not stipulating the characteristics the outcome should have, but stipulating the social procedure that has to be followed. Obviously, it is self-evident that an account of consensus is social. Consensus is always a product of sociality. However, we want to emphasize the parallel of our accounts with *social accounts of objectivity*[8]. Another intuitive way to grapple the difference is to imagine when one calls a certain decision democratic in contemporary political theory: this is not because the decision or outcome of the procedure has certain intrinsic characteristics that make it democratic, but because itself is the result of going through a democratic procedure. Shifting the focus from the product to the epistemic processes resulting in an outcome, also implies that consensus comes in degrees, depending on the extent to which the procedure has been followed, repeated, etc.

The debate about *scientific consensus* then moves to consensus on epistemic procedures, i.e. finding a form of meta-consensus. Thus, the tension between scientific plurality and consensus is not tackled on the simple level, but on the meta-level. This is analogous to how democratic societies deal with value pluralism; the focus is not on getting rid of value pluralism, but on establishing a framework - a meta-consensus - within which pluralism can be dealt with satisfactorily. The meta-consensus can be one that prescribes rational deliberation (in line with models of deliberative democracy), or aggregation (stipulating a procedure for adding up the available views), or agonistic pluralism (developing a procedure or constellation - conflictual consensus - that wants to optimize the *epistemic fecundity* via agonism, cf. [AUTHOR], 2009), etc.

5. Consensus-making and the meta-consensus II: deliberative choreographies, aggregation and contestability.

Let us now return to Solomon's analysis of consensus conferences and the procedures at play. Solomon scrutinizes NIH Consensus Development Conferences, as it is a good example of a social, deliberative epistemic procedure.

Advocates of rational deliberation as most desirable social epistemology have been claiming it reveals bias and presuppositions, corrects errors, generates additional evidence, and more can be accomplished through rational deliberation among two or more individuals than by individuals working alone. For some, for instance, Helen Longino (1990), this rational deliberation and critical interaction are constitutive of scientific objectivity, provided that the interaction respects some norms.[9] The NIH consensus program also works hard to be perceived as objective (cf. Solomon, 2007: 174), although it may not actually achieve objectivity *qua* freedom from all bias:

[8] The most well-known example here is Helen Longino's social account of objectivity (cf. Longino 1990, 2002). See also Douglas (2004) for a good analysis.

[9] We do want to add, though, that for Longino objectivity comes in degrees and she does not consider the "closure" of controversy or critical interaction as the outcome to be obtained – plurality and dissent have their value. This is contrary to most consensus conferences where the "closure" of intellectual controversy and the establishment of a consensus is imperative.

> "The consensus program 'science court' is not designed to be free of biases such as, for example, group dynamics, ordering of speakers, rhetorical force of speakers, peer pressure, chair style, general medical practice biases (e.g. intervention is generally favored over non-intervention), unsystematic evaluation of evidence, the effects of sleep deprivation and conservativeness or radicalism of panel members. The only biases it is designed to eliminate (and it may or may not succeed in doing so) are those of governmental pressures, commercial pressures and biases from one's own prior research in the area." (Solomon, 2007: 169)

Here, Solomon emphasizes via her discussion of consensus conferences that rational deliberation does not deliver what it promises as it is subject to groupthink, suppression of dissent and other biases typical of deliberating groups (also see her 2006). Instead of the procedure of deliberation, Solomon wants to defend the procedure of aggregated judgment, in which members of a group typically do not deliberate with each other, but instead cast their votes or give their views independently. This might even make the NIH consensus conferences obsolete: "A more recent epistemic concern about the idea of consensus development is, wouldn't it be quicker, more timely, and at least as good to do a meta-analysis of the available evidence? Such a formal analysis would have a similar claim to be free from bias." (Solomon, 2007: 169)[10][11] Should the conferences play an epistemic role at all, that is: "Consensus conferences seem to miss the intended window of epistemic opportunity: they typically take place *after* the experts reach consensus." (Solomon, 2007: 170)

But, is the epistemic work done, if one limits the procedure to aggregation? Are we not losing ***epistemic adequacy*** and ***contestability*** (linked to *epistemic responsibility*) out of sight? Returning to the analytical distinction between an academic and an interface consensus, Solomon seems to presuppose that an aggregative procedure establishing academic consensus does all the epistemic work (and establishing an interface consensus could possibly be understood as doing merely policy work). However, the NIH conferences could perform epistemic functions that the aggregation neglects.

a) epistemic adequacy

In her analysis of NIH conferences, Solomon notices that: "The NIH consensus program has never been assessed for the accuracy of outcomes. No-one has ever investigated, for example, whether the outcomes are better – more 'true' or whatever – than those achieved by other methods such as non-neutral panels or formal meta-analysis of evidence." (Solomon, 2007: 174) When she scrutinizes social deliberative epistemic procedures wondering whether they yield the desired kinds of products, she focuses on accuracy of outcomes, not on their adequacy. Adequacy seems to be an afterthought.

[10] A bit further in her paper Solomon correctly describes that meta-analysis of evidence has now been part of the NIH Consensus Conferences preparations. She thus asks whether there is still room for consensus conferences when meta-analyses have already been performed. What Solomon refers to here as meta-analysis is the systematic literature review prepared by the Agency for Healthcare Research and Quality (AHRQ), which is a literature review that tries to identify, appraise, select, and synthesize all high quality research evidence relevant to a particular research question ([AUTHOR], 2012).

[11] See also Sunstein (2006) for a development of this line of criticism, we will not engage with that here.

Let us briefly illustrate the difference between accuracy and adequacy that we have in mind here. Speaking about (a) *accuracy*, we think of the relation with reality, preciseness of the answer given, while (b) *adequacy* points at what the explainee expects of an answer or how it fits with the explainee's epistemic interests. To clarify these criteria and the idea that there often is a trade-off to be made between them, let us compare the answers given to the questions of a consensus conference with maps. A subway map like the one of the Paris Metro is *adequate* for its users because it *accurately* represents specific types of features (e.g. direct train connections between stations, number of stations between two given stations, …) while other features are consciously *less accurately* represented (the exact distances between the stations, the relative geographical orientation of the stations, …). If the latter would be represented more accurately, the map could become less *adequate* for its intended users and a perfectly accurate representation mirroring every detail would be utterly useless. Other maps (e.g. Paris' shopping or tourist attractions maps) require other kinds of information (relating to, e.g., distances, details about street names, house numbers, etc.) in order to be useful – the best trade-off between accuracy and adequacy differs depending on the interests or desiderata at play. Thus, on the one hand, because of different interests or desiderata, it is impossible to make a map that is ideal in all possible situations. On the other hand, not all maps are equally good, as one can make claims of superiority that are bound to specific situations. The same can be said about the answers to the questions that are put central in a consensus conference.

In Solomon's understanding of what NIH conferences actually do, the *accuracy* of outcomes should not be a worry, given she claims that there is already an academic consensus before the conferences take place which seems to make the consensus conference itself obsolete from an epistemic point of view. She sees NIH conferences as the (interface) moment in which the wording or the articulation of this consensus is done in an *adequate way* ("user-friendly"): "Second, formal meta-analysis of research, classified into grades of quality of research, is hardly user-friendly. The NIH consensus statement is written so as to be intelligible not only to primary health care practitioners but to health care administrations and the general public. NIH conferences are not only rhetorical forces; they make the research more widely accessible." (Solomon, 2007: 175) However, this understanding of the relation between accuracy and adequacy, between the academic moment and the interface moment, is too uni-directional, as if adequacy concerns do not touch upon accuracy concerns. Adequacy concerns can be articulated in the interface moment, which then should be fed back into the academic moment in order to trade-off accuracy and adequacy concerns. In that sense the interface moment is not merely a moment of user-friendly dissemination, but also one of critique and contestability of the academic moment (highlighting epistemic interests that have not been addressed).

Solomon defends her claim in relation to the NIH conferences, by drawing on two examples, i.e. the 'Helicobacter Pylori in Peptic Ulcer Disease conference' (1994) and the 'Management of Hepatitis C Consensus Development Conference' (2002). By analyzing both in detail, we can show how the interface moment does highlight epistemic interests that have not been

addressed at the academic moment and thus point at the epistemic role of the interface moment in terms of adequacy[12]:

First, according to Solomon, "the 2002 Consensus Development Conference 'Management of Hepatitis C' 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 something more substantial than merely repeating FDA recommendations, it takes these recommendations to another level and incorporates them in a larger framework. Addressing the adequacy in this case, means answering the above 5 questions, a task the NIH consensus conference set out to do. The multitude of questions were not yet addressed

[12] As is stated on the NIH website, the consensus conference examples we refer to here serve merely an historical and epistemological purpose to defend claims on consensus formation. The actual statements referred to from within any of the reports should thus not be judged for their truth today: "This Archive of Older Conference Statements is provided solely for historical purposes. Due to the cumulative nature of medical research, new knowledge has inevitably accumulated in these subject areas in the time since these statements were prepared. Thus some of the material is likely to be out of date, and at worst simply wrong. The statements may, however, continue to be useful to the research community as a reference for understanding what was known about a topic at a particular point in time, including whether gaps in research identified at the time of each conference have since been filled. It is for this purpose that the conference statements will remain available in this format indefinitely. For reliable, current information on these and other health topics, we recommend consulting the National Institutes of Health's MedlinePlus http://www.nlm.nih.gov/medlineplus/ (NIH, website on the consensus development conference program)".

in academic literature, nor the question what the best therapy might be[13]. Clearly an example of how epistemic interests surpass the academic moment.

Second, 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 prominent clinicians, had reached consensus on the use of antibiotics for peptic ulcers" (Solomon, 2007: 170). The NIH 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). However, this is not the only matter this report investigated. Similar to the 2002 report mentioned above, this report brought together specialists in gastroenterology, surgery, infectious diseases, epidemiology, and pathology, as well as the public, to address multiple questions: (1) What is the causal relationship of H. pylori to upper gastrointestinal disease? (2) How does one diagnose and eradicate H. pylori infection? (3) Does eradication of H. pylori infection benefit the patient with peptic ulcer disease? (4) What is the relationship between H. pylori infection and gastric malignancy? (5) Which H. pylori-infected patients should be treated? (6) What are the most important questions that must be addressed by future research in H. pylori infections? After presentations by experts and discussion by the audience, the consensus panel weighed the evidence and prepared their consensus statement. Among their findings, the panel concluded that: (1) ulcer patients with H. pylori infection require treatment with antimicrobial agents in addition to antisecretory drugs whether on first presentation with the illness or on recurrence; (2) the value of treating nonulcer dyspepsia patients with H. pylori infection remains to be determined; and (3) the interesting relationship between H. pylori infection and gastric cancers requires further exploration (NIH, 1994: 3-4). Solomon's claim that there was a pre-established consensus on the use of antibiotics for peptic ulcers, seems a bit to straightforward, as the answer to the fifth question warrants further research to be conducted, at least for some of the patients intended: There are ample data to support the antimicrobial eradication of H. pylori infection in patients with peptic ulcer disease. All patients with gastric or duodenal ulcers who are infected with H. pylori should be treated with anti-microbials regardless of whether they are suffering from the initial presentation of the disease or from a recurrence. H. pylori-infected peptic ulcer patients who are receiving maintenance treatment with antisecretory agents or who have a history of complicated or refractory disease should also be treated for the infection. The presence of NSAID's, including aspirin, as a contributing factor should not alter the antimicrobial regimen, but whenever possible, these drugs should be discontinued. *However, in asymptomatic H. pylori-infected patients without ulcers, the data are not sufficient to support prophylactic antimicrobial therapy to prevent ulcer disease in the future to reduce the likelihood of developing gastric neoplasia. Also, no convincing data exists to support routine treatment of patients with nonulcer dyspepsia who are infected with H. pylori. Thus, at the present time there is no reason to consider routine detection or treatment of H. pylori infection in the absence of ulcers. Carefully controlled prospective studies are*

[13] This is a critique that transcends this particular report, as most NIH consensus reports address a multitude of questions targeted at the needs and interests of the people (see also the NIH website on the consensus development program).

needed to assess the benefits of treating nonulcer dyspepsia patients with H. pylori infection. It is self-evident that no patient should be treated for H. pylori unless one of the sensitive and specific tests previously discussed demonstrates infection (NIH, 1994: 14, our italics).*"* This answer suggests how for some cases definitive answers need to be postponed as long as prospective studies remain absent. Another example here, is how the question whether eradication of H. pylori infection prevents gastric cancer, can only directly be answered by use of a long and costly study, which was not present. The NIH report therefore suggests an alternative approach to conduct studies looking at the intermediate endpoints that are thought to predict the evolution of malignancy and their response to H. pylori eradication. Also, the report recommends that more epidemiological studies would be needed to define more precisely the subset of H. pylori-infected individuals who will develop gastric cancer (NIH, 1994: 17). Moreover, a causal relationship between H. pylori and peptic ulcer disease is more difficult to establish (as opposed to a causal relationship between H. pylori and chronic superficial gastritis) from the available data in part because of the lack of an animal model and because only a small proportion of individuals harboring the organism develop ulceration (NIH, 1994: 5). All of these are thus obstacles H. pylori research is confronted with in ascribing the correct use of antibiotics. Adequacy concerns, as mentioned above, can thus be articulated in the interface moment, in which participants can advance epistemic interests that have not been addressed, which then should be fed back into the academic moment in order to trade-off accuracy and adequacy concerns. In this case, a requirement for prospective studies, intermediate endpoint studies , epidemiological studies and animal models serve as a feedback loop, where the interface moment, by addressing the question of appropriate medication, gives advice on what research still needs to be conducted to be able to serve the needs of those infected[14]. In that sense the interface moment is not merely a moment of user-friendly dissemination, but also one of critique and contestability of the academic moment.

Perhaps the easiest way to grasp the opportunity the NIH consensus conferences leave for adequacy is in how they value the evaluation of the patient as most central in their discourse. For instance, in the 2000 consensus conference on osteoporosis, it says that: "[…] until there is good evidence to support the cost-effectiveness of routine screening, or the

[14] In other cases, such as the NIH consensus conference on management of hepatitis B (2008), similar demands for further research as to be able to answer the question for treatment appropriately reemerge. The conclusion of this report goes as follows: "[…] Hepatitis B is a major cause of liver disease worldwide, ranking as a substantial cause of cirrhosis and hepatocellular carcinoma. The development and use of a vaccine for hepatitis B virus (HBV) has resulted in a substantial decline in the number of new cases of acute hepatitis B among children, adolescents, and adults in the United States. However, this success has not yet been duplicated worldwide, and both acute and chronic HBV infection continue to represent important global health problems. Seven treatments are currently approved for adult patients with chronic HBV infection in the United States: interferon-a, pegylated interferon-a, lamivudine, adefovir dipivoxil, entecavir, telbivudine, and tenofovir disoproxil fumarate. Interferon-a and lamivudine have been approved for children with HBV infection. *Although available randomized, controlled trials (RCTs) show encouraging short-term results – demonstrating the favorable effect of these agents on such intermediate markers of disease as HBV DNA level, liver enzyme tests, and liver histology – limited rigorous evidence exists demonstrating the effect of these therapies on important long-term clinical outcomes, such as the development of hepatocellular carcinoma or a reduction in deaths. Questions therefore remain about which groups of patients benefit from therapy and at which point in the course of disease this therapy should be initiated* (NIH, 2008: 4, our italics)." The NIH report here tries to meet these questions on patient treatment, that purely on the basis of scientific evidence (academic consensus in Solomon's account), are left unaddressed.

efficacy of early initiation of preventive drugs, an individualized approach is recommended (NIH, 2000: 19)." Taking into account adequacy during consensus conferences will thus influence the content of the academic consensus – which makes the conference more than merely dissemination and choreography.

b) contestability (and *epistemic responsibility*)

Different epistemic interests to be addressed, and related trade-offs between accuracy and adequacy, might result in a variety of outcomes in which the idea of articulating *thé* scientific consensus becomes more nuanced. Adequacy concerns explicated during the interface moment put the plurality of epistemic interests up front. Interface assures contestability and raises questions about the degree of answerability/responsibility of the experts/scientists. Contestability - providing participants the possibility to challenge homogenization, marginalization of epistemic interests, etc. and diminish these tendencies – can be found both on the simple level of finding consensus as well as on the meta-level, i.e. the consensual setting in which the interaction takes place (and on which there is some form of meta-consensus).

An aggregative procedure misses contestability, in the sense that, as Alison Wylie puts it: "[…] well functioning aggregation […] preserves systematic bias as faithfully as it preserves the information and empirically probative insights held by members of a group" (Wylie, 2005: 46)[15]. Moreover, in many occasions, we lack the knowledge, means and time to assess whether the conclusions been put forth are the right ones. When we take one's testimony for granted, we do this often purely based on trust. For this we must have good reasons to trust the other person, in other words, he or she must be deemed trustworthy (Hardwig, 1991: 697-700). The question that springs to mind is whether we can expect of people to trust a judgment that was taken through aggregation, where the possibility exists that the group decision does not correspond with any individual opinion defended in the group. When, for instance, a group is polarized on an issue and the decision is but an average not defended by any of the people involved. A dubious policy that no one actually endorses, would thus be the result of a strongly divided group. Aggregative procedure taking is but a quasi-transparent process: when there is no opportunity for debate and no arguments are given in favor or against a certain take on things, it remains difficult for the general audience to acquire on which grounds a decision was taken ([AUTHOR], under review). This but obscures the opportunity for contestability, as part of the process remains in the dark for most of the people involved. An interface moment could allow for contestability as it offers opportunity both to (1) question bias, as well as, (2) acquire insight into how certain positions and decisions came to be. The NIH general procedure allows for both moments of contestability: Because they want to provide an independent look on the topic, they opt for a type of 'court model', instating an unbiased panel. The review-process itself is structured as follows: First, there is an in-depth presentation of evidence to the panel. This includes a

[15] As an example Wylie investigates research concerning prerequisites on gender. She mentions there being research in which professors were to judge a fictive candidate for a job on the basis of a resume. When the candidate was named Mark Miller, he was hired by two thirds of the professors. However, when the identical resume made mention of Karen Miller, the candidate was but hired by less than half of the professors (Wylie, 2006: 46).

systematic literature review prepared by the Agency 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 (NIH, website). It is only but through the procedure that moments of contestability become possible.

Examples of contestability concerning meta-consensus, i.e. the procedure, can be found, for instance, in relation to the NIH Consensus program, as we learn from Solomon's analysis (2011: 244): "So we see particular concerns about "objectivity" of panel members, fairness of chairs, time pressure and late night sessions, balanced assessment of the evidence, and so forth. The NIH CDC Program has been evaluated on a number of occasions: internal review in 1980, a University of Michigan Study in 1982, a Rand Corporation review in 1989 an IOM study in 1990, and most recently by a NIH working group in 1999 [Perry and Kalberer, 1980; Wortman *et al.*, 1982; Kanouse *et al*., 1989; IOM, 1990a; Leshner *et al.*, 1999] Concerns have regularly been expressed about panel selection to ensure 'balance and objectivity,' speaker selection that represents the range of work on the topic, representation of patient perspectives and more careful and systematic assessment of the quality and quantity of scientific evidence. Concerns have also been expressed about the time pressure to produce a statement in less than three days, and especially the lack of time for reflection or gathering of further information. Such concerns have in fact been behind some changes over time in the NIH CDC Program, and also behind the creation of different procedures at other kinds of consensus conferences, in both the national and the international scene."

Kristina Rolin (2009) stipulates the epistemic role outsiders to particular scientific communities can play. She argues that an epistemically responsible scientist has a duty to respond to outside criticism in certain circumstances insofar as it includes an appropriate challenge to her views. A meta-consensus taking contestability into account differs from both aggregated judgment – in supporting dynamical, diachronic interaction – and rational deliberation – avoiding groupthink via contestation. Obviously, contestability comes in degrees and is present to a greater or lesser extent in the existing formats for consensus-making.

Having pointed at the epistemic importance of considerations of adequacy and contestability, we can conclude that the role of the consensus conference is not merely choreographic or 'just' a matter of policy. We highlighted clear epistemic roles, both in explicating adequacy and in assuring contestability – it is an epistemic role which Solomon misses in her critique; "If the topic is not a scientific topic but is, instead, a matter of policy, the use of a consensus conference can be appropriate. The scientific community can tolerate— even celebrate — research disagreements. Policy decisions usually require the *joint action* of individuals, groups and nations. A well-negotiated consensus is widely thought of as the ideal foundation for joint action. So those European Consensus Conferences which

consider questions of policy are more appropriately designed to attain their goals than, for example, the NIH CDC Program." (Solomon, 2011: 247-248)

6. Modifying the public understanding of scientific consensus.

Establishing scientific consensus is highly valued by philosophers[16], scientists and the public. The emergence of a scientific consensus replacing competing accounts is often interpreted

[16] There are also other philosophers that have dealt with the value of pluralism and consensus. A similar example as to the way we dealt with both terms in this paper, is the recent work by Martin Carrier (2012). I here briefly present his account and how it relates to ours: According to Carrier, in addition to pluralism at the level of theories and value-commitments alike, scientific research is also characterized by a joint striving for consensus which he traces back to a shared epistemic attitude. This attitude manifests itself, e.g., in the willingness of scientists to subject their claims to empirical scrutiny and to respect rational argument. This shared epistemic attitude is embodied in rules adopted by the scientific community concerning general principles of dealing with knowledge claims. His contention is that pluralism and consensus formation can be brought into harmony by placing them at different levels of consideration: at the level of scientific reasoning and at the level of social conventions regarding how to deal with claims put forward within the scientific community. That is to say, the belligerent pluralism constitutive of essential parts of scientific rationality is curbed or pacified by a joint striving for consensus. Pluralism remains temporary and transient; it comes to an end eventually and gives way to consensus. (Carrier, 2012: 16) In particular, Carrier argues that we regularly observe that a period of strife and confrontation is followed by the emergence of a consensus. When the dust has settled, a convergence toward widely shared views obtains in the scientific community. Further, according to him, this consensus is achieved without external pressure. Some sort of internal mechanism appears to ensure that the plurality of contrasting and heterogeneous approaches gives way to a generally accepted state of research. But it remains to be clarified by which process the scientific community settles on a certain hypothesis or theory eventually. One option Carrier considers, is that the large majority of empirical achievements and epistemic values point in the same direction. That is, although values such as broad scope, coherence with background knowledge, and accuracy are conceptually distinct, it may happen that they all favor one particular account. Miriam Solomon subscribes to such a view, according to which consensus in science is rare and without epistemic significance . Consensus is generated by the fortuitous convergence of a variety of factors of a diverse nature, empirical or non-empirical, cognitive or emotional, rational or social. (Solomon, 2001: 11, 99-120) However, the eventual formation of a consensus is typical in science and distinctive of it. The repeated well-considered intellectual withdrawal that goes along with the emergence of a scientific consensus is a hallmark of scientific debates. Political and religious strife tends to continue indefinitely and is only stopped by the weariness of the fighting parties. Yet scientific controversies usually come to an end on substantive grounds. According to Carrier, this phenomenon of giving in deliberately is a characteristic of the scientific enterprise by which it distinguishes itself from other intellectual fields. (Carrier, 2012: 17-18) The epistemic attitude is, according to Carrier, codified in the form of procedural rules of the scientific community. The epistemic attitude does not address the process of assessing hypotheses directly; it rather concerns procedures for debating such assessments. This attitude finds its expression, in particular, in commitments like attending to dissenting views and empirical problems, taking up criticism, and granting intellectual authority on substantive grounds alone . Further, these rules are essentially social: they address how to deal with nonconformist understandings and opposing approaches. Such procedural rules for addressing substantive diversity are suitable for constraining antagonistic beliefs and to drive them toward a common position. Keep in mind that they are also epistemic in that following their advice is conducive to reliability or truth, that is, epistemic goals of science. (Carrier, 2012: 19) Such procedural rules are intertwined with values, but values of a different sort, namely social values of the scientific community. These rules address how to deal with dissenting views and opposing approaches and display a clear epistemic bearing. They are epistemic values of a particular sort. Such social values of the scientific community can be expected to be more stable than the prevalent epistemic commitments. The latter values may change as a result of interacting with nature and of attempting to cut nature at the joints. Yet the former values distinguish preferred patterns of social interaction in examining knowledge claims and thus remain unaffected by changes in the demands of what a good explanation or an excellent theory ought to accomplish. (Carrier, 2012: 17) Carrier, in his analysis, clearly focuses on academic consensus and omits to see the comparison to interface consensus. According to

as a proof of scientific progress and a marker of truth; ideally all scientific inquiry and debate would result in a consensus. Finding scientific consensus is then understood as a proof for the validity of a theory and – indirectly – of the public policy based on the consensus theory.

The back side of the coin is that the lack of scientific consensus often is used to undermine or criticize science and the public policy based on it (e.g., former US President Bush on climate change). When scientists agree, their results are taken more seriously than when they disagree, even though such an agreement or consensus might hinder scientific progress because of critical, heterodox theories not being taken seriously (e.g., the theory of continental drift was accepted by geologists only after 50 years of rejection, and the theory of helicobacter pylori as the cause of stomach ulcers, was at first widely rejected by the medical community).

These observations might question scientific consensus as an ideal or as the goal of inquiry and marker of truth; enforcing 'consensus' might be dangerous or not desirable, hence the importance of scrutinizing carefully what is actually going on in establishing scientific consensus; i.e. which moments one can distinguish, the epistemic and non-epistemic value of those moments as well as the different procedural set-ups – just like we have been doing above in distinguishing the academic, interface and meta-consensus. Communicating this variety of formats to the public, helps qualifying the actual span of scientific consensus making and the oracle like features it might sometimes have. Our reasoning is in line with Inmaculada de Melo-Martin and Kristen Intemann's recent paper, where they show, by analyzing conflicts on GMO's and climate change, that 'focusing on dissent as a problematic activity sends the message to policy-makers and the public that any dissent undermines scientific knowledge.' (de Melo-Martin & Intemann, 2013: 232-233) Encouraging and providing mechanisms of dissent can also be important to reassuring the public that the consensus view has undergone rigorous scrutiny. Events such as "climate-gate" reinforce the public perception that climate scientists are resistant to criticism and have a "bunker mentality" (Grundman, 2012).

Pointing at some of the undesirable effects too high expectations about consensus-making might have, we refer to Churchill's famous dictum: "Democracy is the worst form of government, except for all those other forms that have been tried from time to time." (from a House of Commons speech on Nov. 11, 1947). Similarly, science and consensus-making mechanisms create fallible knowledge.

These insights taken together do not foreclose scientific consensus-making, but stress the need for a mentality change in society about what scientific research consists of and what it can entail for policy making. As such, ""in a scientific community, different individuals can weigh evidence in different manners through the use of different standards. In the best

---

him, the epistemic attitude is an attitude that differentiates the scientific community from the broader society, as science allows to eventually reach consensus on any given topic. However, as we have shown above, persistent disagreement is also a characteristic of science and not merely of science policy. Moreover, Carrier endorses Solomon's view on consensus conferences, of scientific consensus being rare and without epistemic significance, which we refuted in the previous section of this chapter. Although Carrier manages to pinpoint the benefit of a procedural account, he omits to see how a procedural account of consensus can be opened up to include policy science.

case[17], science puts forward a robust consensus based on a research process that allows continued scrutiny, re-examination, and revision."(Oreskes, 2004: 369-370)[18]

7. Conclusions

Through our analysis we have argued in favor of the following claims:

a) Consensus-making as discussed by philosophers of science, should be aware of the difference in aiming for an *academic consensus* or an *interface consensus* in science and society. A broader understanding of the different structures and functions of consensus-making helps us to see more nuances.
b) The difficulties of achieving consensus understood as an unanimous outcome – a seldom-attained ideal stipulated by a range of criteria – should make us shift to a procedural approach in which the emphasis is not so much on establishing the consensus, but dealing with plurality in a consensual way, i.e., framed within a meta-consensus that agrees on how to disagree.
c) Taking into account epistemic adequacy and contestability as important characteristics of sound epistemic processes, we questioned Solomon's account of consensus conferences as just being moments of deliberative choreographies, repeating work that has already been done by the aggregation of expert opinions.
d) Solomon's idea of (what is the valuable) scientific consensus as, basically, an idea of *academic consensus* as we described it, holds on to an ideal which we also encounter among the public in general. According to us, it is too high an ideal, that eventually can be used against science; it is therefore more recommendable to think in terms of degrees of consensus and in adequacy for the public addressed while keeping an eye on the variety of consensus-making formats.

[17] A sentiment shared by other researchers: "The idea of consensus in science does not imply the fact that all the scientists have internalized and agreed upon the truth of the statements that make up a certain consensus. In this sense, what is called a 'scientific consensus' in the literature can be the product of compromise, negotiation, and only under special circumstances a truly consensual resolution" (Martini, 2011: 152).

[18] This quote can read, as Miller would probably do, saying that this is not always the case, still thinking that only consensus that has certain properties should be trusted. (based on personal communication with Boaz Miller) However, we read "in the best case" as pointing out the fact that in many cases no robust consensus will be found, but nevertheless consensus seeking might be optimalized to (1) acquire the best result possible, (2) prevent forced consensus being established, and (3) allow trust to be generated.

References

Baartmans, T. & Kosolosky, L. (under review). Groepsbeslissingen: kwaliteit, autoriteit en vertrouwen*. Tijdschrift voor Filosofie.*

Beatty, J. (2006) 'Masking Disagreement among Scientific Experts' *Episteme* 3: 52-67.

Beatty, J. and A. Moore (2010) 'Should we Aim for Consensus?' *Episteme* 7(3): 198-214

Carrier, M. (2012) 'Values and Objectivity in Science: Value-Ladenness, Pluralism and the Epistemic Attitude', *Science and Education*, DOI 10.1007/s11191-012-9481-5.

De Melo-Martin, I. & Intemann, K. (2013) 'Scientific dissent and public policy: Is targeting dissent a reasonable way to protect sound policy decisions?', *European Molecular Biology Organization reports*, 14(3): 231-235.

Douglas, H. (2004) 'The irreducible complexity of objectivity' *Synthese* 138(3): 453-473.

Dryzek, J. and S. Niemeyer (2006) 'Reconciling Pluralism and Consensus as Political Ideals.' *American Journal of Political Science* **50**(3), 634-649.

Gilbert, M. (1987) 'Modeling Collective Belief.' *Synthese* 73 (1): 185–204.

Grundmann, R. (2012). 'The legacy of climategate: revitalizing or undermining climate science and policy?', WIREs Clim Change, doi: 10.1002/wcc.166

Hardwig, J. (1991). The Role of Trust in Knowledge. *Journal of Philosophy*, 88(12): 693-708.

Kosolosky, L. (2012). The intended window of epistemic opportunity: A comment on Miriam Solomon, In: Bart Van Kerkhove, Thierry Libert, Geert Vanpaemel & Pierre Marage, (eds.), Logic, Philosophy and History of Science in Belgium II., Koninklijke Vlaamse Academie van België, Brussel, 2012.

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

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

Mansbridge, J. et al. (2010) 'The Place of Self-Interest and the Role of Power in Deliberative Democracy.' *The Journal of Political Philosophy* 18(1): 64-100.

Martini, C. (2011). *Consensus and Disagreement in Small Committees*. Phd-thesis.

Miller, B. (2013). When is consensus knowledge based? Synthese, 190(7): 1293-1316.

National Institute of Health: Official Site, Web Site: Available online: http://www.nih.gov/

National Institute of Health. Website on the Consensus Development Conference Program. Available online: http://www.consensus.nih.gov

National Institute of Health (1994). Helicobacter pylori in peptic ulcer disease. NIH Cons Statement. February 7-9, 12(1): 1-23. Available online: http://consensus.nih.gov/1994/1994HelicobacterPyloriUlcer094PDF.pdf

National Institute of Health (2000). NIH Consensus Statement: Osteoporosis Prevention, Diagnosis, and Therapy. March 27-29, 17(1): 1-45. Available online: http://consensus.nih.gov/2000/2000Osteoporosis111PDF.pdf

National Institute of Health (2002). Management of hepatitis C: 2002. NIH Consensus Statements. June 10-12, 19(3): 1-46. Available online: http://consensus.nih.gov/2002/2002HepatitisC2002116PDF.pdf

National Institute of Health (2008). NIH Consensus Development Conference Statement on Management of Hepatitis B. October 20-22, 25(2): 1-29. Available online: http://consensus.nih.gov/2008/hepbstatement.pdf

Oreskes, N. (2004). Science and public policy: what's proof got to do with it? Environmental Science & Policy, 7: 369-383.

Rolin, K. (2009) 'Scientific Knowledge: A Stakeholder Theory.' In: Van Bouwel, J. (ed.) (2009), *The Social Sciences and Democracy*, Basingstoke: Palgrave Macmillan, pp. 95- 119.

Solomon, M. (2006). '*Groupthink* vs. *The Wisdom of the Crowds:* The social epistemology of deliberation and dissent.' *Southern Journal of Philosophy*, volume 44, Supplement: 28–42.

Solomon, M. (2007) 'The social epistemology of NIH consensus conferences.' In: H. Kincaid and J. McKitrick (eds) *Establishing medical reality: Methodological and metaphysical issues in philosophy of medicine.* Dordrecht: Springer.

Solomon, M. (2011) 'Group Judgment and the Medical Consensus Conference.' In: D. Gabbay and J. Woods (eds.) *Handbook of The Philosophy of Science: Philosophy of Medicine*. San Diego: North Holland, pp. 239-254.

Van Bouwel, J. (2009) 'The Problem with(out) Consensus. The Scientific Consensus, Deliberative Democracy and Agonistic Pluralism.' In: J. Van Bouwel (ed.) *The Social Sciences and Democracy.* Basingstoke: Palgrave Macmillan, pp. 121-142.

Wylie, A. (2006). Socially Naturalized Norms of Epistemic Rationality: Aggregation and Deliberation. *The Southern Journal of Philosophy*, volume 44, Supplement: 43-48.