Consensus and Collaboration in Contemporary Scientific Practice

Michelle Trang Pham

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Andrea Woody, Chair

Carole J. Lee

Conor Mayo-Wilson

Program Authorized to Offer Degree:

Philosophy

University of Washington

Abstract

Consensus and Collaboration in Contemporary Scientific Practice

Michelle Trang Pham

Chair of the Supervisory Committee:
Andrea Woody
Philosophy

I argue that there is no single, monolithic conception of 'consensus' in scientific practice. I distinguish between two categories of aggregative and collaborative consensus. Collaborative consensus, which is produced by interdependent groups, has not been adequately explored in the context of contemporary scientific practice. A central contribution in my dissertation is the proposal of normative conditions for well-functioning collaboration. The trustworthiness of collaboratively-produced consensus positions should include consideration about whether they meet these proposed normative conditions. I also use the distinctions and analyses discussed to assess the public controversy regarding the IPCC's consensus position on anthropogenic climate change, looking at the arguments of skeptics and a prominent defender, Naomi Oreskes. Both Oreskes and skeptics, I argue, wrongly invoke aggregative consensus to assess the IPCC's position. However, the group is a collaboration, and their consensus position cannot be assessed according to standards of aggregation. I instead assess the IPCC using the proposed normative conditions for well-functioning collaboration. I show that they generally satisfy them, and I then provide a different response to skeptics for why there are good social reasons to trust the IPCC.

Table of Contents

List of Figures

Acknowledgments

This dissertation was a collaborative effort with many contributions from others, especially my committee members. We may not share the same beliefs on everything, but I think we were motivated by some broader collective goals. First, thank you to my GSR, Bruce Hevly, for his comments on the prospectus and the final draft of the dissertation. Thanks to Carole Lee for her inquisitive questions and comments on all of the Chapters, but especially discussions on Longino and co-authorship. She also helped me shape the initial project, working with me on my literature review. I also want to acknowledge all the other ways she supported me as a mentor. Thanks to Conor Mayo-Wilson for his contributions, especially his multiple sets of comments for Chapter 3. He encouraged me to include consideration about the potential impact of my work in the real-world. I also benefitted from discussions in his Social Epistemology seminar. Finally, I would like to express gratitude to my Chair, Andrea Woody. Many of the arguments in this dissertation were crystalized during our dissertation meetings. This dissertation benefitted from her tough questions and suggestions. I am grateful to Andrea for her enthusiasm about this project, for giving me freedom to explore different literatures, and for letting me know when it was time to move on. I admire her care for her students and her expectations for intellectual rigor.

This project also benefitted from many conversations with both graduate students and faculty in the philosophy of science community. In addition, I benefitted from the conversations in Leah Ceccarelli's Rhetoric of Science seminar. I'd like to thank Leah for her comments on Chapter 3.

I would like to thank my CNT Neuroethics collaborators. Sara Goering and Eran Klein are both invaluable mentors, and I learned much about collaborative practice working with them. As well, over the years, I received much support and advice from Sara Goering. I also would like to thank my fellow postdoc, graduate, and undergraduate Neuroethics collaborators for their weekly conversations – I learned a great deal from them.

Outside of philosophy, I am lucky to always be supported by family. Thanks to Jon, and Cindy & Mike for supporting my intellectual journey over the years. Finally, I am grateful to my husband, Steve, as well as Chemi, for encouragement, especially during the final phase of dissertation writing.

Introduction

Many pressing contemporary public policy decisions made by governmental bodies rely on scientific expertise as an important source of information. However, during the past few decades scientific recommendations have come under attack by parties who either disagree or have conflicting interests (see Oreskes and Conway 2010). These attacks contribute to rampant misinformation about scientific knowledge, and they undermine the legitimacy of experts among various lay communities. Misinformation has always been present but the ways in which people recently access information online, curated by big tech corporations using algorithms designed to maximize profits, is quite different from what we've seen before.[1] The ubiquity of information online, both true and false, means readers can find seemingly legitimate support for "both sides" of almost any given issue. This gives the appearance that "both sides" are genuine competing alternatives. For example, that climate change is both increasing the prevalence of extreme weather patterns and droughts, *and* that it is a ploy for more government control on people's liberties – basically, a hoax. Some other public controversies that have ignited debate and interest online include: the efficacy of vaccines for children (whether the MMR vaccine causes autism) and whether evolutionary theory and creationism should be taught side-by-side in high school science curriculum. If enough individuals and communities of various lay publics are either misinformed or simply do not trust scientific experts (or both), then it is difficult for governmental bodies to craft and implement effective public policies based on the most current scientific knowledge.[2] Central to many of these public controversies between defenders of science and skeptics is a focus on whether there is a genuine consensus among scientific experts. Consensus of experts has become one standard by which to gauge the legitimacy of an expert scientific group's position.

Despite this very public focus on the status of expert scientific consensus, some philosophers (de Melo Martin and Intemann 2018; Goldenberg 2019) have recently suggested that the lay public's mistrust in science is not rooted in misinformation, "poor scientific literacy," and/or poor communication about the consensus of experts. Rather, it is in relation to science's questionable institutional practices, such as the increasing commercialization of science, researchers' conflicts of interest, and scientific misconduct. I should also add that science's history is marred with unjust and inhumane research, such as the Eugenics Movement during the early 19th century, the Tuskegee Syphilis Experiment from 1932-

[1] Some tech corporations are beginning to pay attention to concerns about the proliferation of misinformation, but it is unclear how they will be accountable and to whom. Moreover, it is unclear whether their acknowledgement of the problem is genuine or rooted in PR concerns.

[2] Matters are worse when members of democratic governments themselves believe in false information and (intentionally or not) promote conspiracy theories. Unfortunately, some do (see Siegler 2020).

1972, and the medicalization of homosexuality, to name a few. So, perhaps efforts to educate the public about science, and further emphasizing that scientists have a consensus position regarding a public policy issue to correct the ubiquity of misinformation will not help gain the public's trust. While I agree with these authors that scientists ought to build trustworthy relationships in part by revising science's problematic institutional practices, I think that if a group of experts have a consensus position, then such consensus matters for public policy decisions. However, supporters of science would need to offer an explanation to lay communities about why the consensus of experts matters beyond appeals to it being a mark of scientific authority (Goodwin 2009). I suggest one alternative: that scientists appeal to the consensus of experts as a mark of epistemic trustworthiness by pointing out that their community's procedures and practices promote a well-functioning group. Yet very little has been said about what counts as consensus that emerges from a well-functioning group, which can be confusing for lay persons who are unsure of why they should care if experts agree or disagree. This dissertation articulates some answers to fill this gap by offering an analysis of consensus as a central social feature of scientific practice.

We better understand scientific consensus, I argue, when we recognize that there are distinct types rather than one monolithic 'consensus'. I broadly distinguish between aggregative and collaborative consensus.[3] Aggregative consensus involves a procedure where individual views are counted or tallied, and the total "votes" meet some minimum threshold such as a majority for there to be a consensus position. We see aggregative consensus take place in survey studies that aim to capture the aggregate view of a specific group. Differently, collaborative consensus involves cohesively organized teams that interact and partake in joint action to produce a group position. Procedures for determining a group position can vary but members might, for example, agree to let some position stand as the group's view to align with some set of broader goals. Often, collaboratively-produced consensus positions involve negotiation among members who divide epistemic labor and who may not necessarily share the same expertise. We see instances of collaborative consensus emerge from common collaborative practices, like co-authorship and large-scale research collaborations. Importantly, aggregative and collaborative consensus have different standards of assessment. When gauging whether a consensus position is legitimate we need to first identify what type it is, and not mistakenly use the standards of one to assess the other. While the distinction between aggregative and collaborative consensus is central to this dissertation, attention will primarily be paid to the details of collaborative consensus. Let me explain why.

During the last few decades, collaboration has become a central component of scientific research (Morrison et al, 2003). For example, co-authored publications are now a norm, making up 64.7% of

[3] As I will discuss in Chapter 1, there are also consensus models that integrate both aggregative and collaborative components, so there are cases in which they overlap. Still, this general distinction is useful for making sense of certain fraught public debates about expertise with important policy implications, such as the climate change case (see Chapter 3).

worldwide science and engineering publications in 2016 (NSF 2018). The NSF also reports that international collaborations have increased for worldwide publications from 16.7% to 21.7% between 2006 and 2016.[4] For the US, publications that include co-authors affiliated with international institutions went up from 25.2% in 2006 to 37% in 2016. A variety of reasons have been offered for why scientists collaborate, which include: defraying risk, expanding one's social network in a broader community, and sharing labor in contexts where one cannot complete a project alone for various reasons. An important consequence of collaboration, especially for interdisciplinary and large-scale groups, is that no single individual is in a position to know everything about the project. As Boyer-Kassem et al (2018) point out about climate science research, for example: "No single researcher has the empirical knowledge, mathematical training, and programming abilities to construct the models of clouds, precipitation, glacier movement, and so on, that are employed in such climate studies" (2018, xi). This has important implications for how we assess the legitimacy and trustworthiness of joint-products that collaborations offer. Though social epistemologists and philosophers of science have recently taken an interest in scientific collaboration (see Andersen and Wagenknecht 2013; Boyer-Kassem et al 2018; Wagenknecht 2015; Wylie 2015), few have made an explicit connection between collaboration and consensus.

A central goal of this dissertation is to offer a more detailed analysis of collaborative consensus and to propose normative criteria or guidelines for collaborations that produce group positions. I take the view that group positions that emerge during collaboration are basically consensus positions (e.g., stated views in co-authored papers, conference presentations, and joint statements to the public). In trying to make sense of what we mean by consensus for groups in science, and determining what its overall function is in practice, looking at how collaborations produce group views provides fruitful direction. A focus on articulating and clarifying collaborative consensus is significant because so little has been said about a social phenomenon that is central to significant scientific products, such as co-authored publications, joint public statements, and policy recommendations. Furthermore, some of the fraught public controversies involving scientific expertise rest on (implicit) debates about whether the consensus position of a *collaborative group* is legitimate (see Chapter 3).

In looking at significant types of collaborative practice, including research collaboration and co-authorship, I argue that what makes scientific collaboration distinct are the relations of dependence (i.e., epistemic dependence and social roles) between group members. These dependence relations reveal that collaborative groups are often structured or organized hierarchically, since members occupy different social roles with unequal distribution of authority based on factors like efficient division of labor, expertise, experience or career stage, and prestige. I propose four normative criteria for well-functioning collaboration (fair credit attribution, respectful participation, group leader accountability,

[4] NSF 2018, Chapter 5. Available at: https://www.nsf.gov/statistics/2018/nsb20181/report/sections/academic-research-and-development/highlights.

and external grievance reporting) that takes into account this hierarchical structure exhibited by many collaborations.

I also make use of the public controversy about anthropogenic climate change in the US as a central case study to apply and illustrate the importance of the philosophical concepts and distinctions discussed. In particular, I show how the distinction between aggregative and collaborative consensus can help clarify the debate between skeptics and defenders (i.e., Naomi Oreskes) of climate science. I argue that both proponents wrongly invoke aggregative consensus to discuss the IPCC's position on climate change, even though the group has produced a collaborative consensus position. I then use the proposed normative criteria to assess the IPCC collaboration and argue that though there is room for improvement, the group generally meets the proposed conditions for well-functioning collaboration. This, I argue, is why the IPCC's collaboratively produced consensus position should be trusted. It is also why the skeptics are wrong to insist there is no (aggregative) consensus among climate scientists of the IPCC by appealing to dissenting scientists. The IPCC's collaborative consensus position cannot be assessed according to the aggregative model because of the dependence relations members share with each other. No single individual is in a position to assess all there is to know about the comprehensive report, which involves a sophisticated division of cognitive labor with experts who possess highly specialized knowledge. Rather, an assessment of the IPCC's consensus position should be based on whether the collaboration itself satisfies normative conditions for well-functioning collaboration – and it does.

This dissertation also makes an original contribution to the philosophy of science and social epistemology of science literature. In the last few decades, philosophers of science have taken what's referred to as "the social turn" by focusing on key social components of science or the processes of interaction and communication among practitioners, especially on how to maintain productive dissent and diverse perspectives (see Longino 1990, 2002; Solomon 2001; Wylie 2012).[5] But few discussions grapple with the import of consensus as a central social feature of science. One noteworthy exception, Miriam Solomon (2001; 2006) has argued that consensus may be epistemically bad for science because it can emerge as a result of groupthink, where group members concede to a view because of pressures to conform. While Solomon's concerns are legitimate, I maintain that it does not mean that we give up on consensus altogether. I think these concerns are mitigated when we get clear on the sort of consensus under discussion, and we locate the social practices where these types of consensus take place. One reason so little progress has been made on the role of consensus in the philosophical literature has to do with its elusiveness and a lack of connection to science's practices. Overall, my analysis of consensus involving the resources of conceptual literature, real-world case studies, and scientific practice will demonstrate why the social turn ought to include consensus as an important social feature of scientific inquiry. In the next section, I provide a brief summary of each chapter.

[5] Some other social aspects of science that have been discussed recently include bias in the peer review process (Lee 2013; Lee and Schunn 2010), how to properly disseminate information in the age of big data (Leonelli 2014), and credit (Bruner and O'Connor 2018; Lee 2020).

Dissertation Preview

1 Aggregative vs. Collaboration Consensus? In Chapter 1, I distinguish between two broad categories of consensus relevant to scientific practice: aggregative and collaborative consensus. I provide a taxonomy of consensus models in the philosophy of science and social epistemology of science literature under this distinction, as well as models with mixed methods that fall in-between. My aim, unlike some of the debates that have emerged, is not to take a position on whether aggregative or collaborative models of consensus are better or worse in comparison to each other but rather to discuss their merits and drawbacks in terms of usefulness. I argue aggregative and collaborative consensus models serve different purposes. Aggregative consensus models help pool information from individual group members; whereas for collaborative consensus models, group members interact and work together in their pursuit of a joint endeavor. Furthermore, I argue that the standards by which we evaluate these two categories of consensus ought to be different (a point articulated in Chapters 3 and 5). In addition, while this distinction is central to the discussions that occur in later chapters, a primary focus will be on getting clear on what makes collaborative consensus distinct, and what counts as well-functioning collaboration.

2 Resources for Analyzing Collaborative Consensus: Evaluating Gilbert's Joint Commitment Model. Chapter 2 defends a prominent collaborative consensus model in the social epistemology of science literature, Margaret Gilbert's joint commitment model, against a set of criticisms from Melinda Fagan. First, I argue that supporters of the joint commitment model can adequately respond to Fagan's criticisms. Second, I argue that even though Gilbert's model does not capture the social processes in science by which joint commitments emerge, this is not a particularly pressing problem for the model's application to scientific contexts. The joint commitment model was meant to be broad in scope, capturing an everyday social phenomenon in a variety of human social contexts. It is the philosopher of science's and/or social epistemologist of science's job to articulate the relevant practices in science in which joint commitments occur, and to identify the social processes involved in them. Gilbert's model offers a useful conceptual framework to consider consensus in scientific practice, but it is by no means an exhaustive account. In addition, other models (e.g., Tuomela's positional group belief, and Beatty and Moore's deliberative acceptance) that focus on different social features, often viewed as being in competition with Gilbert's, should be viewed as helpful conceptual models too. They are all envisioning particular types of circumstances, none of which alone are sufficient for capturing scientific practice. In this sense, I do not find it useful to strongly argue in favor of one model, while rejecting others. My approach will be to use these models to help conceptualize consensus that emerges from collaborative practices.

3 Considering Consensus in the Climate Change Controversy. Chapter 3 uses the distinction between aggregative and collaborative consensus from Chapter 1 to make sense of the debate about the status of climate change between a prominent climate science defender, Naomi Oreskes, and climate change skeptics. I point out that both Oreskes and the skeptics invoke aggregative consensus to assess

the status of the IPCC's broad anthropogenic climate change claim. This, I argue, is a mistake. In particular, the IPCC's consensus position on anthropogenic climate change is a collaboratively produced one that aligns well with understanding of collaborative consensus. So while I commend Oreskes' efforts to confront climate skeptics, I do not think her argumentative strategy helps clarify the confusing debate for various lay publics. In the next two chapters, I articulate what makes collaborative groups that produce consensus position distinct, and what normative conditions for well-functioning collaboration ought to look like.

4 Scientific Collaboration: Epistemic Dependence and Social Roles. Chapter 4 provides an analysis of scientific collaboration by focusing on the social dynamics of collaborative practice. What is missing from many of the collaborative consensus models (i.e., those discussed in Chapter 1) is an analysis of the social dynamics in research contexts from which collaborative consensus is formed. Many of those models, including Gilbert's, only capture the end point for which the group view is established with little insight regarding how collaborative consensus positions emerge in scientific practice. I fill in this gap by looking at two important collaborative practices in science: research collaboration and co-authorship. I argue that the dependence relations members share with one another have a significant role in how collaborative groups function – specifically, relations of epistemic dependence and hierarchical social roles assigned to group members. Assuming this focus on the significance of dependence relations is correct, I then consider Longino's influential account of scientific inquiry, "critical contextual empiricism," as potentially offering normative conditions to assess collaboration. However, in my analysis I conclude that Longino's normative conditions are not appropriately applicable because they do not consider authority structures often exhibited in scientific collaboration. I thus offer my own normative conditions for collaboration in the next chapter.

5 Normative Criteria for Collaboration. Chapter 5 proposes normative criteria for collaboration: fair credit attribution, respectful participation, group leader accountability, and external grievance reporting. These conditions take into account the dependence relations discussed in Chapter 4. I argue that while epistemic dependence and social roles are crucial to productive collaboration, they can also lead to problematic social dynamics, such as those related to groupthink (in line with concerns voiced earlier by Solomon (2001; 2006)) and abuse of authority. The normative conditions I propose aim to assuage problematic social dynamics. I then assess the IPCC collaboration according to the proposed normative criteria and argue that the IPCC meets conditions for well-functioning collaboration. This, I argue, provides an answer to the climate science skeptic. First, the IPCC's consensus position should not be evaluated according to standards of aggregation. The IPCC group is a collaboration. Second, I show the IPCC collaboration meets conditions for well-functioning collaboration, which gives us reasons to trust the group's procedures and practices that helped produced their consensus position. I end the discussion with some concluding remarks.

Bibliography

Andersen, Hanne, and Susan Wagenknecht. 2013. “Epistemic Dependence in Interdisciplinary Groups.” *Synthese* 190: 1881–98.

Boyer-Kassem, Thomas, Conor Mayo-Wilson, and Michael Weisberg, eds. 2018. *Scientific Collaboration and Collective Knowledge*. New York: Oxford University Press.

Bruner, Justin, and Cailin O’Connor. 2018. “Power, Bargaining, and Collaboration.” In *Scientific Collaboration and Collective Knowledge*, edited by Thomas Boyer-Kassem, Conor Mayo-Wilson, and Michael Weisberg, 135–57. New York: Oxford University Press.

Goldenberg, Maya J. 2019. “Vaccines, Values and Science.” *Canadian Medical Association Journal* 191 (14): E397–98.

Goodwin, Jean. 2009. “The Authority of the IPCC First Assessment Report and the Manufacture of Consensus.” *Working Draft*.

Lee, Carole J. forthcoming. “The Reference Class Problem for Credit Valuation in Science.” *Philosophy of Science*.

———. 2015. “Commesuration Bias in Peer Review” 82 (5): 1271–83.

Lee, Carole J., and Christian D. Schunn. 2010. “Social Biases and Solutions for Procedural Objectivity.” *Hypatia* 26 (2): 352-373.

Leonelli, Sabina. 2014. “Data Interpretation in the Digital Age.” *Perspectives on Science* 22 (3): 297–417.

Melo-Martín, Inmaculada de, and Kristen Intemann. 2018. *The Fight Against Doubt*. Oxford University Press.

Morrison, Philip, Gill Dobbie, and Fiona J. McDonald. 2003. “Research Collaboration Among University Scientists.” *Higher Education Research & Development* 22 (3): 275–96.

National Science Foundation. 2018. “Science and Engineering Indicators 2018.” https://www.nsf.gov/statistics/2018/nsb20181/report/sections/academic-research-and-development/highlights.

Oreskes, Naomi, and Erik M. Conway. 2010. *Merchants of Doubt*. New York: Bloomsbury.

Siegler, Kirk. 2020. “In Idaho, Far-Right Republications Defy Coronavirus Health Restrictions.” NPR. April 13, 2020. https://www.npr.org/2020/04/13/832961674/in-idaho-far-right-republicans-defy-coronavirus-health-restrictions.

Solomon, Miriam. 2001. *Social Empiricism*. Cambridge, MA: MIT Press.

———. 2006. “Groupthink versus the Wisdom of Crowds: The Social Epistemology of Deliberation and Dissent.” *The Southern Journal of Philosophy*, no. 44 Supplement: 28–42.

Wagenknecht, Susan. 2015. “Facing the Incompleteness of Trust: Managing Dependence in Scientific Practice.” *Social Epistemology* 29 (2): 160–84.

Wylie, Alison. 2012. “Feminist Philosophy of Science: Standpoint Matters.” *Proceedings and Addresses of the APA* 86 (2): 46–76.

———. 2015. “A Plurality of Pluralisms: Collaborative Practice in Archaeology.” In *Objectivity in Science*, edited by Flavia Padovani, Alan Richardson, and Jonathan Y. Tsou, 189–210. Boston Studies in the Philosophy and History of Science 310.

1 Aggregative vs. Collaborative Consensus?

As I finish this dissertation in the midst of the COVID-19 pandemic, the number of deaths from the virus has surpassed 100,000 in the United States and over 350,000 globally. The US's death rate is the highest in the world, and this is perhaps a surprise because the US has some of the best medical science experts and produces some of the best medical research globally. Recently, protests have been organized in state capitals against stay-at-home orders, which were crafted to prevent the spread of coronavirus. Some of these protestors, which include anti-vaxxers, claim that the scientific evidence offered exaggerates the effects of the virus and that the orders violate constitutional rights to freedom. Some even went as far to say that the pandemic is just another hoax to remove the sitting president (Schwartz et al 2020). To make matters worse, the current administration has sent mixed messages to the public about the severity of the pandemic by either contradicting or even rejecting guidelines offered by the Centers for Disease Control (CDC), which is comprised of an expert body.[6] These sorts of responses involving falsehoods, misinformation, and conspiracy theories about the scientific evidence behind COVID-19, against the CDC's guidelines, however, are not unique to the ongoing public health crisis. They are part of a brewing undercurrent over many decades that is now in full view: doubt and distrust in scientific expertise.

One of the ways in which skeptics have contributed to the public's distrust of scientific experts is by sewing doubt about science through the proliferation of false information (e.g., the tobacco industry's years of public denial and "research" that rejected the connection between smoking and lung cancer (see Oreskes and Conway 2010)). More specifically, one strategy is to reject the consensus position of a group of experts by suggesting that the position is either false because there are actually dissenting opinions on the issue not properly represented or that the consensus is manufactured. In the case of COVID-19, the CDC's current consensus position on hydroxychloroquine as a potential treatment, crafted by an expert body, is that it currently is not an effective option. The CDC's position, articulated by the NIH states:

> The Panel **recommends against** the use of chloroquine or hydroxychloroquine for the treatment of COVID-19, except in a clinical trial. (NIH COVID-19 Treatment Guidelines, 2020; original emphasis)[7]

The current president, however, has repeatedly touted the benefits of this drug. His evidence for its effectiveness is that doctors and nurses are taking it as a preventative measure:

[6] This, of course, is not to say that the CDC has been perfect in their communications with the public.
[7] The WHO also recently halted testing of hydroxychloroquine in their large study searching for potential treatment for COVID-19 due to an observational report in the Lancet noting that patients receiving the drug were dying at higher rates than other coronavirus patients (Beaubien 2020).

> "[A] lot of people are taking — a lot of front-line workers are taking hydroxychloroquine."
>
> "But you look at front-line workers. You look at doctors and nurses. A lot of them are taking it. As a preventative." (*Politico* 2020)

His reasoning, indirectly against the CDC (since the CDC also advises the president), is that hydroxychloroquine is effective because frontline medical workers, who are an experienced group of experts dealing with coronavirus every day, are taking it. This contradiction with the CDC's recommendation suggests the CDC is wrong. In addition, the president has also claimed that the only reason why hydroxychloroquine is not getting more positive attention is because, in his words, "[W]e have this crazy whistleblower, this fake whistleblower [Rick Bright, the ousted director of the Biomedical Advanced Research and Development Authority] get out and try to knock it" (*Politico* 2020). So not only are doctors and nurses taking the drug as a preventative measure, it turns out a fraud with dishonest intentions is trying to prevent its use. For those who support the president, they may thus be led to believe that the CDC's public position on hydroxychloroquine is false because (a) there are dissenting opinions (i.e., doctors and nurses seem to think it's effective since they are taking it) and (b) that the negativity surrounding the drug is manufactured (i.e., Rich Bright is a fake whistleblower). This same strategy of undermining the recommended consensus position of an expert body can be seen in many other public debates, including climate change and MMR vaccines.

In many ways, sociologists of science Harry Collins and Robert Evans (2002) anticipated the dysfunctional and concerning public discourse on scientific expertise taking place today. They recognized nearly 20 years ago the need for a normative account of expertise about issues of real consequence for diverse publics. They proposed "third wave science studies" which takes the descriptive lessons from the sociology of scientific knowledge (SSK) programme seriously, especially that science is a social activity. But in addition Collins and Evans theorize expertise in terms of the relationships among experts who are able to interact and contribute to the science under investigation, not just those who are technically "certified" as scientists (e.g., under this framework Cumbrian sheep farmers in the Chernobyl case are considered experts) (2002, 261). Their analysis offers a more fined-grained and inclusive understanding of expertise; indeed, they helpfully point out that many scientists do not have sufficient training or knowledge to conduct analysis in scientific fields in which they are not versed. However, on the issue of expert consensus, they are less helpful. Collins and Evans write: "[D]ecisions of public concern have to be made according to a timetable established within the political sphere, not the scientific or technical sphere; the decisions have to be made *before* the scientific dust has settled, because the pace of politics is faster than the pace of scientific consensus formation" (2002, 241). Collins and Evans are right that public policy decisions requiring immediate action cannot wait for "absolute consensus formation."[8] However, public policy makers do still need to rely on expert scientific bodies for recommendations.

[8] It is interesting that Collins and Evans mention 'consensus' 17 times in their discussion paper, but they never quite say what it is. The closest we get to a definition is the phrase "absolute consensus formation," which I take to signal consensus that is considered to be something like established knowledge.

Such recommendations are of course not the same as "absolute consensus formation," since they may change with more available information. Still, these recommendations are the scientific bodies' group position, established through consensus. If we want to help improve public discourse on matters related to scientific expertise, then in addition to a normative theory of expertise regarding who counts as an expert, a normative theory of consensus that offers guidelines for what counts as trustworthy expert consensus recommendations is also needed.

During the same year Collins and Evans published their discussion paper, an influential monograph from the philosophy of science, offering a normative theory of scientific inquiry, came out: Helen Longino's *The* Fate *of Knowledge* (2002). This monograph, like Collins and Evan's discussion paper, takes seriously the SSK's point that science is an inherently social activity.[9] Indeed, Longino makes use of science's sociality as a strength: critical interaction can help mitigate bias and give attention to diverse perspectives. Longino proposes "critical contextual empiricism" which offers normative conditions for critical interaction among diverse scientific communities to achieve objective knowledge acquisition. However, despite the influence and significance of Longino's work, especially in the philosophy of science community, it stays within the bounds of scientific inquiry and offers few resources for considering expert scientific consensus. So while both Collins and Evans, and Longino make important contributions to science studies, broadly construed, we are in dire need of a normative theory of expert scientific consensus, one that could help improve the public discourse on scientific expertise. The current global health epidemic, which has been fraught with confusion, mixed messaging, and outright attacks on individual and groups of experts demonstrates this.

Similar to the authors discussed above, in this dissertation I work with the assumption that science is an inherently social activity, and that many of science's social features contribute to scientific knowledge acquisition.[10] In the course of this dissertation I will show that one of these crucial social features is consensus. Here I have in mind joint endeavors that result in group positions such as co-authored papers and publicly declared group stances. And like many other social features of science, there can be better and worse consensus positions depending on how they are achieved. For the purposes of this project, I think we need to first clarify what we mean by 'consensus' in the context of scientific inquiry. To this end, I will distinguish between two broad categories of consensus – aggregative and collaborative consensus, and discuss specific models offered in the literature that fall under these two categories.[11] I argue that aggregative and collaborative consensus are not mutually exclusive; each should be understood as appropriate for particular goals or purposes. The methods used for each category can also overlap in important ways. Rather than frame the conversation as one about having to choose

[9] Also see Kitcher (2001).
[10] Readers can see both Collins and Evans (2002) and Longino (2002) for arguments that support this claim.
[11] These, of course, are not the only categories of consensus. For my purposes though, I think these categories will help improve the analysis over the course of this discussion.

between aggregative and collaborative consensus to make sense of consensus in science, I propose taking a purpose-oriented approach that focuses on particular circumstances in practice.

This chapter will be organized as follows. In section I, I discuss aggregative consensus, focusing on Miriam Solomon's "normatively appropriate consensus" model, with some discussion of Steve Fuller's "accidental consensus." I also discuss some strengths, as well as some weaknesses of these models when determining consensus. In section II, I discuss collaborative consensus, focusing on Margaret Gilbert's "joint commitment" model and Raimo Tuomela's "positional group belief." I also discuss some of their strengths and potential weaknesses. In section III, I propose framing aggregative and collaborative consensus as being useful for making sense of consensus in particular practical contexts, rather than as being in tension. I give examples of contexts for when aggregative and collaborative consensus models can be usefully employed. I also highlight how features of each model are sometimes employed simultaneously by one group, so they can be mutually supportive. Finally, I briefly discuss why the remainder of this dissertation will focus largely on collaborative consensus models.

I Aggregative Consensus

Aggregative consensus generally tallies the views or beliefs of a community's members to establish a consensus position. For example, if a group is trying to reach consensus on, say, whether a drug is effective for treating a particular disease, each member will make an individual judgment about the drug's effectiveness. The totality of those individual judgments are then aggregated to form the group position. The threshold for what constitutes an appropriate aggregation result will depend on what the group decides the standard to be. They could, for example, require unanimous consensus where the aggregative result is acceptable if and only if each member makes the same judgment. Alternatively, they might require a majority where the aggregative result is acceptable when more than half of group members make the same judgment or perhaps a calculated average. Two features are worth highlighting. One is that each member's judgment is included. The other feature is that aggregative consensus allows for individual members to provide independent judgment about the issue under consideration, sometimes without any consultation from others depending on the model.

With the general framework of aggregative consensus just highlighted, one might think that Steve Fuller's (1986) discussion of the "accidental consensus" model is a good exemplar of aggregative consensus.[12] Fuller discusses historical cases of multiple or simultaneous discoveries, where scientists make roughly the same discovery within a close time frame without communication – so no discussion or deliberation. For example, the independent discovery of the energy conservation principle by Carnot, Joule, Helmholtz, Mayer, and Rumford in the 1830s since, according to Fuller, they had "little or nothing

[12] Note that Fuller does not endorse accidental consensus but discusses it to reveal the elusiveness of consensus.

to do with each other's" discovery (1986, 109). One might think simultaneous discoveries are an indicator of independent judgment or "objectivity" because it seems more likely that such a discovery is accurate, and not a fluke, when multiple people come to the same conclusion on their own. Notice that if accidental consensus is an apt exemplar of an aggregative consensus, then by implication what constitutes appropriate consensus will be *unanimous* consensus where all group members share the same belief.[13] Also note though that on this model the historian (or a historically-minded scholar) makes the judgment about whether the beliefs of the individuals tallied are competent ones, and so they determine who is part of the community when assessing whether there was consensus.

However, though accidental consensus seems to possess many of the general features of aggregative consensus that I mentioned above, it is not a good exemplar of aggregative consensus. First, Fuller notes that "[…] the 'multiplicity' of this discovery [e.g., the energy conservation principles] may simply be - as Kuhn [...] has suggested - the result of historians [...] finding consensus by stating the principle abstractly enough" (1986, 109). Multiple discoveries may look like objective knowledge but only because the historian has described the relevant phenomenon in an abstracted way that makes it convenient for the numerous instances of "discovery" to fit under that same umbrella description. Without access to their private reasons for justifying the principle we can't be sure if their versions of the energy conservation principle are identical and/or relevantly similar to each other's.[14] Second, the purpose of this project is to analyze scientific consensus in contemporary scientific practice with real-world impact. Setting aside the issue of whether in historical cases, scientists who are doing completely independent research are actually referring to the same terms or concepts, a condition of *complete* independence is unrealistic and perhaps even detrimental when considering contemporary scientific practice. The reason is that contemporary scientific practice is hardly ever conducted completely independently from others; instead, scientists and researchers publish in venues that their peers have access to, they share and read each other's work, and they attend the same conferences. An important strength of current scientific practices is the myriad ways in which researchers are able to share their work and receive feedback. For these reasons, accidental consensus does not serve as a helpful exemplar of aggregative consensus. That said, I do think it is helpful to discuss a model that, in the most extreme form, captures independent judgment, and reveal why, despite claims about complete "objectivity," it may not be fruitful for my purposes. Let's move on to Miriam Solomon's proposal, which I do take to be an exemplar of aggregative consensus.

Normatively Appropriate Consensus

Solomon argues that, at least on epistemological grounds, consensus is not the end point of scientific inquiry, and, in fact, it should be a relatively rare phenomenon in science. The presence of

[13] I will have some points of discussion about unanimous consensus shortly.
[14] Ultimately, Fuller's position on scientific consensus is that it is an elusive concept.

dissent in a scientific community should be considered a strength because it keeps a variety of perspectives on the table that premature consensus would otherwise rule out while prizing others. Premature consensus potentially forecloses fruitful avenues of inquiry. But in those relatively rare cases of consensus, what are the conditions by which it might be warranted? Solomon calls such cases "normatively appropriate consensus" where the amount of dissent in a given community "goes to zero" (2006, 37). Focusing on scientific theories, Solomon argues normatively appropriate consensus must meet three conditions for the absence of dissent to be acceptable (2006, 37). First, the consensus view must have *all* of the empirical successes: the "observation[al], predictive, retrodictive, experimental, and explanatory or technological" successes (Solomon 2001, 27). Second, the theory must have all of the empirical decision vectors – empirical aspects that contribute to scientists' judgment that a theory is empirically successful, such as salience of data and availability of data. Third, non-empirical decision vectors must over time support the consensus theory – these are factors independent of a theory's empirical success but contribute to scientists' preference for it such as pride, simplicity, ideology, and competitiveness. The independent judgments of individual scientists are aggregated to figure whether a theory captures all of the available empirical successes and empirical decision vectors, and whether the non-empirical decision vectors over time support it. The individual judgments of scientists, however, is not what is of central importance.[15] Solomon writes: "What matters is not how individual scientists reason – *it's* not *the thought that counts* – but what the aggregate community of scientists does" (2001, 135; original emphasis).

One historical example Solomon discusses is the consensus on plate tectonics (2001, 102-109).[16] She writes: "Consensus was complete when plate tectonics had all the empirical successes (those that permanentism, contractionism and drift had, and more besides)—but not because any individual or individuals saw that it had all the successes" (2001, 109). Note that even though Solomon's articulation and discussion of normatively appropriate consensus in her book *Social Empiricism* (2001) focuses on scientific theories, it has broad applications for determining group knowledge generally, as she discusses in her more recent paper in relation to the groupthink phenomenon (see 2006).[17] In that paper, Solomon highlights an example from James Surowiecki (2004) about the missing *Scorpion* U.S. submarine:

> The Navy had only the vaguest of information about where it might be: somewhere in a circle twenty miles in diameter and thousands of feet deep. Craven, a naval officer, assembled a diverse group of experts and asked each to individually calculate their best prediction of the submarine's path and find location. There was no discussion or consultation of these experts with each other. John Craven combined the calculations using Bayes Theorem and came up with an

[15] Solomon writes: "normative judgments are not made of the thoughts and decisions of individual scientists" (2001, 120).
[16] See Oreskes's (2008) exchange with Solomon (2008) for some interesting points of disagreement regarding this historical case study.
[17] More on groupthink shortly.

aggregate prediction. It turned out to be just 220 yards from the discovered location of the *Scorpion*. (Solomon 2006, 35)[18]

Here the group view is determined by aggregating each individual expert's independent calculation using the framework of Bayes Theorem to make a prediction, which turned out to be accurate. In both of these examples, notice that no discussion or deliberation takes place. Solomon writes: "*[A]ggregation without deliberation preserves information*" because it prevents undue influence from one's peers and seniors." (2006, 36).

Some Strengths and Weaknesses

So far, I have highlighted some general features of aggregative consensus, and discussed Solomon's normatively appropriate consensus as one exemplar. I want to now discuss some strengths, as well as weaknesses, of aggregative consensus. One strength is that the absence of discussion and deliberation means social phenomena like groupthink, where individuals consent to a view because of pressure from peers or authoritative group members, are less likely to take place. This is precisely why Solomon advocates for establishing consensus through aggregation methods, rather than discussion and deliberation. She argues that discussion and deliberation may actually lead groups to worse consensus positions. Solomon writes: "[...] there is little doubt that factors such as peer pressure, pressures from authorities, pressure to reach consensus, and the salience of particularly vocal group members who may anchor decisions lead groups into making poor decisions" (2006, 32). Aggregative consensus, on the other hand, seems less vulnerable to social phenomena like groupthink, and may thus be preferable because it does not involve social interaction.

The other strength I want to highlight is empirical evidence showing that aggregating the judgments of individuals can lead to more accurate information. For example, Francis Galton's (Surowiecki 2004) well-known experiment aggregated 800 participants' estimate of an ox's weight, generating a result which was only one pound less than the actual weight – 1197 to 1198 pounds, respectively.[19] Christian List and Phillip Petit (2011) call this "information pooling," in which they write: when "pooling diverse information of multiple individuals, a group can outperform all of its members at truth tracking" (2011, 86). A group tracks the truth when they accurately judge the truth or falsity of some proposition.[20] What better way to rely on a consensus position then if everyone's judgments, when aggregated, provides more accurate information? It seems promising that members of a group independently make judgements, and together those judgments are more trustworthy than any particular individual's.

---

[18] See Surowiecki (2004, xx-xxi).
[19] Solomon also discusses this case (2006, 34).
[20] Their notion of truth-tracking follows Robert Nozick's definition where two conditionals must hold: "[…] if 'p' were true, the agent would judge that p" and "[…] if 'p' were false, the agent would judge that not p" (2011, 82).

Aggregative consensus also has some weaknesses. I will briefly mention two. One issue pertains to the independence feature in which members make judgments on their own. It turns out that this is a rather high standard to achieve since influencing factors can manifest in many other contexts than just those involving discussion and deliberation. Though Solomon's concerns about the emergence of groupthink during discussion and deliberation is legitimate, she does not consider the fact that contemporary scientists indirectly interact with their peers by reading research in peer-reviewed journals, and more directly by participating at conferences.[21] Moreover, when they publish in journals, which most are required to do, it is the norm that they are expected to be versed in and also properly cite the relevant literature. Their papers are also given feedback from anonymous reviewers (usually other scientists with the relevant expertise), as well as the journal editor(s), which they are required to address to the journal editor(s)' satisfaction. There are many ways for scientists to be influenced by the judgments of their peers that do not involve face-to-face discussion and deliberation. In this sense, contemporary scientists' judgments are rarely completely independent from the influence of others. It is unlikely that we would get uninhibited and independent judgments from scientists simply by employing aggregative methods to form consensus.

This does not mean that scientific knowledge is doomed because practitioners are influenced by each other's ideas. The value of information pooling mentioned earlier tells us that collective knowledge can help yield more accurate information. It is also a hallmark of contemporary scientific practice that practitioners rely on others' work and collaborate on projects they cannot complete alone. So while I take Solomon's concerns about groupthink seriously, the difficulties she raises with group discussion and deliberation should not result in complete rejection. Rather, we should find ways to mitigate the potential emergence of groupthink. As Alison Wylie puts it: "Rather than reject deliberative processes in favor of aggregative techniques, the norms of epistemic rationality inspired by these processes should incorporate a detailed, empirically grounded understanding of the conditions under which group deliberation can work well, and of the conditions under which it manifestly fails" (2006, 47).

Another problem with aggregative consensus pertains to results from social epistemology research showing that aggregation functions can be subject to complications. One aspect that does not get much attention is the different ways judgment aggregation could be organized to determine a group's consensus position. Christian List and Phillip Pettit (2011, 48) use the term "aggregation function" to refer to different mapping relations between individual judgments and the group judgment. Using their work, let me briefly highlight two aggregation functions to illustrate different ways an aggregative model may determine a group's consensus position. The "unanimity rule" model, for example, requires unanimous agreement among group members' judgment about 'p' – so each group member must individually judge that 'p'. Unanimous agreement about 'p' is necessary and sufficient for the group position that 'p' (2011, 88). Differently, there is the "majority voting" model, where the majority

[21] Also see Wylie's (2006) paper for a response showing that aggregative methods can be just as problematic as deliberative ones.

constitutes more than half of group members. There is a group position when the individual judgments of group members regarding 'p' constitute more than a majority.[22] Both of these aggregation functions have problems.

Employing a unanimity rule to achieve consensus is practically challenging since the group will not have a group consensus position unless they have unanimous agreement. List and Pettit's results show that as the group size increases, the group's ability to accurately judge 'p' when it is true goes toward zero - so harder to achieve (2011, 88-89). In practical circumstances, such as policy decision-making contexts, this is not a useful strategy to achieve consensus, especially if groups are pressed for time, which might instead require a negotiated position. Further, unanimity may be nearly impossible to achieve if the group is comprised of individuals with diverse or different background expertise. In such contexts, they may not be able to have a group judgment at all. What this suggests is that unanimity rule as a standard to measure the appropriateness of an aggregative consensus position may neither be feasible or useful.[23]

Employing the majority voting model, which groups often do, faces a paradox. Again, List and Pettit's (2002; 2011) work shows that aggregating the judgments of individuals can sometimes lead to counterintuitive results where the group's aggregated judgments on a set of premises does not match up with their judgment on the conclusion. This is known as the "doctrinal paradox" from social choice theory.[24] Take for example a three-member court that needs to rule on a breach of contract case. The court needs to make judgments on three related propositions:

P1 The defendant was contractually obliged not to do a certain action.

P2 The defendant did that action.

∴ The defendant is liable for breach of contract. (2011, 44)

The judges individually made judgments on the related propositions as follows.

	Obligations?	Action?	Liable?
Judge 1	True	True	True
Judge 2	True	False	False
Judge 3	False	True	False
Majority	**True**	**True**	**False**

(2011, 44)

Judge 1 believes both premises are true, so the defendant is liable. Judge 2 does not believe premise 2, so the defendant is not liable. And judge 3 does not believe premise 1, so the defendant is not liable. Aggregating each judges' ruling on the conclusion, the majority of judges rule that the defendant is not

[22] They also discuss the "dictatorship model," which uses one fixed individual's judgment about 'p', so the group's position is whatever the dictator's judgment about 'p'. Other aggregation functions include supermajority (2/3s or more), inverse dictatorship, and constant rules (2011, 49).

[23] See Chapter 3 on how climate skeptics use unanimity as the only standard to measure trustworthy expert consensus on climate change.

[24] List and Pettit's more general formulation of this problem is the "discursive dilemma" (2011, 45).

liable. However, notice that if we aggregate their judgments for each of the premises, the majority of judges believe both premises to be true. But even though the judges believe the premises to be true, they also believe the conclusion to be *false*. This is an inconsistent and paradoxical result since one would assume that if the premises are true, the conclusion must also be true. More generally, the doctrinal paradox identifies a problem with majoritarian aggregation models: groups make inconsistent judgments about related propositions depending on whether decision-making is organized as 'premise-based' or 'conclusion-based'. Ideally, one would assume that the relation between the premises and the conclusion are consistent.

These challenges against the unanimity rule and majority vote models, raised by List and Pettit, have implications for consensus models in science that employ aggregation methods.[25] First, particular aggregation functions face unique challenges. Unanimity rule fairs badly if the group is large, and/or the group may end up without a group judgment since it is practically difficult to achieve. Majority rule is subject to an "impossibility result" between the premises and the conclusion of a related set of propositions. Second, given that there are different ways to aggregate judgments, the aggregated consensus position may be different depending on the aggregation function chosen.[26] The appropriate aggregative function will depend on the group's particular circumstances, including their aims or goals.

More generally, while aggregative consensus models *may* avoid biasing social phenomena like groupthink in some contexts by eliminating discussion and deliberation, they too face challenges. List and Pettit's results should allay over-confidence in the "objectivity" of aggregation methods over more interactive or collaborative ones. My aim though is not to reject aggregative methods but to highlight the complexity of aggregation methods.

II Collaborative Consensus

In this section, I discuss the other category of consensus: collaborative consensus. Similar to the last section, I start with a general description of collaborative consensus, and then highlight models that serve as exemplars of collaborative consensus, focusing on Margaret Gilbert's joint commitment model and Raimo Tuomela's positional group belief. I will then discuss some of their strengths and weakness.

Collaborative consensus is achieved through social interaction and members are interdependent in their collaborative effort. The sorts of groups that achieve collaborative consensus jointly hold a commitment to shared goals, which are common knowledge among group members. This commitment

[25] Solomon (2006) does acknowledges the significance of List and Pettit's discussion of the doctrinal paradox.
[26] List and Pettit argue that the organizational structure of the group can contribute to the group's success or failure.

can be accepted explicitly or tacitly.[27] Importantly, group members work *together* as a unit in their commitment to achieve shared goals. In this sense, collaborative consensus involves interaction, such as discussion and deliberation, so these groups may be highly cohesive. The group consensus position is not achieved simply by tallying the views of individual members to meet some threshold. Rather, it is achieved through acceptance; that is, the willingness of each member to let some position stand as the group's, in service of the *group's* aims. A consequence of the collaborative, interactive, and interdependent nature of groups that work to achieve collaborative consensus is that normative constraints are imposed on group members when they are acting in their capacity *as a group member*. As a group member they cannot, for example, behave in ways that undermine the group's shared goals. In fact, there is an expectation by all that given each member's initial commitment to the group's aims that they behave in ways that help the group successfully realize them. Let's now discuss some collaborative consensus models.

The Joint Commitment Model

Gilbert's joint commitment model is perhaps the most widely discussed model to date that fruitfully captures collaborative consensus (see Beatty 2006; Fagan 2011,2012; Rolin 2008; Schmidt 1994; Wray 2001).[28] The account is defined as follows:

> (1) A group G [accepts] *p* if and only if members of G jointly accept that *p*.
>
> (2) Members of a group G jointly accept that *p* if and only if it is common knowledge in G that the individual members of G have openly expressed a conditional commitment jointly to accept that *p* together with the other members of G. (1996, 204-205)

By "jointly accept," Gilbert means that members are collectively willing to let *p* stand as the group's position. They are "committed as a body in a certain way" (1996, 204). *P* here can stand for a variety of things, such as an action, a claim, or a theory. In addition, *p* is common knowledge among group members, so members are aware that the group accepts *p*. The acceptance of *p* requires that members of G act and behave in ways that not stifle but uphold and promote it. To give a concrete example, John Beatty (2006) uses this model to discuss how during the atomic age a group of geneticists were commissioned to provide a consensus regarding the maximum permissible radiation dosage. The geneticists jointly recommended 10 *roentgens* despite deep internal disagreements. They each agreed

[27] Explicit acceptance of a commitment to a shared goal may occur through social interaction like discussion, deliberation, or negotiation. Newly formed groups like a research team receiving a large body of funding, for example, may decide on their shared goals through explicit acceptance from group members during the planning stages. Tacit acceptance of a commitment to a shared goal may occur when members join a group with pre-existing norms. For example, a graduate student or a postdoc researcher joining a research group with already well-defined goals will tacitly accept those goals. Alternatively, if the group dynamic is such that authority is allocated to one or a sub-group of members to make decisions on behalf of the group, then the remaining members would tacitly accept decisions by those in positions of authority.

[28] Another model is Bratman's (1999) "shared cooperative activity."

to let the position stand as the group's view to fulfill the goal of completing the commissioned task. They also aimed to present a unified consensus position to the public, as expert scientists, to deter the Atomic Energy Commission – an outside interest group – from claiming that radiation effects were negligible. The perception of a unified front would help preserve scientific authority over outside interest groups. The joint commitment model can make sense of how there can be a group view, even though the geneticists disagreed on how to determine the threshold for radiation exposure: by letting the position stand as the group's. In this sense, a group's collaboratively produced consensus position does not simply reflect the beliefs of the individual group members. It emerges as a result of the group's collaborative effort. This is important because members can agree to let a position stand as the group's for different reasons.[29] In this case, the geneticists actually had deep disagreements with each other on how to calculate the appropriate radiation threshold, but in the end they were able to achieve a consensus position for the sake of their joint goal, even when they personally believed there were too many variables to make an accurate and objective calculation.

I should note that under the joint commitment model, group members can accept the group position, *and* personally believe otherwise. There has been much discussion about the difference between belief and acceptance but for this project's purposes it is not of central importance (see Wray 2001). Practically speaking, no philosopher who pays attention to scientific practice would deny that scientists make concessions by operating on certain agreed upon assumptions to promote particular aims, both individual and jointly formed. They may believe it, not believe it, be indifferent, or perhaps believe for reasons other than the ones offered by the group. Of importance is that the joint commitment model can capture how individuals make practical concessions and agree for the sake of a joint goal.[30] This is especially helpful when considering collaborative groups where, different from Beatty's example, the members do not share the same expertise but actually rely on each other's different expertise. In these cases, members must rely on each other for matters they do not know, and so some of the claims they accept will be in deference to the relevant expert in the group.[31]

Positional Group Belief

Another collaborative consensus model I want to briefly mention is Tuomela's (1992) positional group beliefs. Similar to the joint commitment model, it emphasizes that the group position is common knowledge among members. In addition, the model similarly focuses on 'acceptance' where acceptance of something means one will henceforth operate on it as an assumption. What makes this model unique though is Tuomela's articulation of how group members can have different roles and capacities, with

[29] Also see Staley's (2007) discussion of how 450 researchers at the Collider Detector Fermilab Collaboration (CDF) jointly announced evidence for discovery of the top quark.
[30] My aim in this chapter is to simply provide a mapping of the conceptual space for various ways in which consensus may be achieved. In the next chapter though, I will focus attention solely on Gilbert's joint commitment model, and some of the criticisms raised by Fagan (2011; 2012).
[31] See Chapters 3, 4, and 5.

normative constraints imposed on these roles. Two types of role are specified. "Operative agents" are those in positions of authority who make decisions on behalf of the group. "Non-operative agents" do not make such decisions, but they participate through tacit acceptance. The types of actions agents can pursue on behalf of the group are constrained by the group's conventions pertaining to assigned roles, both formal and informal. Here is how Tuomela defines positional group belief:

> G *believes* that p in the social and normative circumstances C if and only if in C there are operative members A1, … , Am of G in respective positions P1, … , such that:
>
> (1') the agents A1, … , Am when they are performing their social tasks in their positions P1, , … , Pm and due to exercising the relevant authority system of G, (intentionally) jointly accept that p; and because of this exercise of authority system, they ought to continue to accept and positionally believe it;
>
> (2') there is a mutual belief among the operative members A1, … , Am to the effect that (1');
>
> (3') because of (1'), the (full-fledged and adequately informed) non-operative members of G tend tacitly to accept – or at least ought to accept – p, as members of G; and
>
> (4') there is a mutual belief in G to the effect that (3'). (1992, 295-296)

The right "social and normative circumstances" are those where members act according to their assigned position in the group (1992, 288). So even though operative agents' "actions, goals, and beliefs are attributed to groups" they do not have free reign to act because their position is constrained by established conventions (1992, 288). Finally, the relevant "authority system" determines the rule-tasks associated with positions in the group, and for the group generally, in a manner that promotes the group view. For example, an authority system might require that sometimes non-operative members completely transfer their authority to operative members, to act on the group's behalf. Some examples of authority systems include laws and statutes or informal group agreements (1992, 288).

Even though Tuomela's model shares many similarities with Gilbert's joint acceptance model, his emphasis on the different roles and obligations of group members is an important point for the purposes of this dissertation. The reason is that collaborations in science are often comprised of group members occupying different sorts of roles with differential degrees of authority. We see this in co-authorship practices and more broadly in research labs both big and small. This explicit acknowledgment of differential degrees of authority in Tuomela's model will be conceptually useful later on when discussing normative conditions for groups that produce collaborative consensus.[32]

Some Strengths and Weaknesses

Let me now discuss some strengths and weaknesses of these collaborative consensus models. An important strength is that they capture the social and interdependent nature of groups that produce collaborative consensus. Group members work together as a unit to realize their shared goals, and their

[32] See Chapter 4 and 5.

collaboratively produced consensus position reflects this. These models help show conceptually how a collaborative group's consensus position is not a summation of the beliefs of individual members but rather a result of social interaction involving discussion, deliberation, negotiation. Furthermore, the organizational structure of groups that produce collaborative consensus is one where members are relationally connected in important ways. Aggregative consensus models do not capture this relational component between group members; rather, they treat members as separate individuals each with the same capacity to make judgments about some particular issue. One of the significant ways in which group members involved in collaboration are relationally connected pertains to the normative obligations associated with their role in the group. For example, the range of acceptable actions or behaviors are constrained by the shared goal(s) that members have committed to uphold and realize. Similarly, these models are useful in group decision-making contexts because once the group makes a commitment to a shared goal that decision would foreground certain decisions and/or actions while precluding others.

Still, one might wonder: why is it important that these models capture the social and interdependent nature of groups that produce collaborative consensus? One reason is that, practically speaking, collaborative consensus is a central feature of contemporary scientific practice as evidenced by the proliferation of joint authorship and public consensus statements put out by scientific bodies such as the NIH and IPCC to inform public policy. As I mentioned in the Introduction of this dissertation, scientific collaboration has steadily risen in the last few decades. Co-authorship, for example, is now the norm and not the exception. Scientists and researchers are working together for a variety of reasons, which include the need to rely on others' expertise that they do have and the ability to divide cognitive labor in ways that would allow them to pursue projects they would not be able to complete on their own. The significance and proliferation of collaboration in contemporary scientific practice has also meant that now more than ever we are increasingly seeing joint consensus positions produced as a result of collaboration in the form of published products like articles and public recommendations or guidelines. Gilbert's and Tuomela's models offer fruitful conceptual resources to capture these collaboratively produced consensus positions. Aggregative consensus models cannot account for the relational aspects of collaborative groups such as divisions of cognitive labor, differential degrees of authority, obligations and deference to others on matters one does not know. Moreover, philosophical discussions that attempt to make sense of consensus in scientific practice that do not account for collaboratively produced consensus miss out on a significant aspect of actual practice.

However, like aggregative consensus models, collaborative consensus models are subject to challenges too. I will briefly mention two. One problem pertains to whether the group's consensus position was established as a result of genuine consent and not coercion, such as the groupthink phenomenon mentioned earlier.[33] Collaborative groups can be vulnerable to negative and toxic social dynamics because of their organizational structure. Even though the close-knit and interdependent

[33] See Chapter 5 for a detailed discussion of groupthink and collaborative consensus.

nature of these groups is integral to the success of collaborations, it also has the potential to promote unhealthy social dynamics where members are pressured and manipulated by others into doing things they otherwise would not. There is, for example, an inequality of authority or power among members depending on which social role one occupies in the group (e.g., PI, postdoc, graduate student researcher). A graduate student researcher may be afraid to speak-up or decline a request to consent to the consensus position if a senior person insisted that they do so for the sake of expediency. So how can it be ensured that collaboratively produced consensus is genuinely formed rather than corrupted through undue influence? Social psychologist Irving Janis (1972) has proposed solutions to mitigate phenomena like groupthink, such as further dividing group members into subgroups to encourage members to voice their opinions in low stakes settings.[34] At this point however there is not enough empirical evidence to say whether organizational strategies like these work. In Solomon's words: "[...] it is important to show that the suggestions can be implemented and that they lead to the desired results. The experimental work is not there yet" (2006, 33). In general, the concern is that collaborative consensus models cannot ensure that a quality group consensus position has been made.

A second problem pertains to the obligations of group members in their capacity as a member of the group when there is disagreement or a change of heart. Recall that once members have accepted the group position, they are expected to act and behave in ways that promote and uphold it. However, what if, for example, a member eventually is unwilling to work with the group's consensus position, even though they originally consented. Are they ever justified in openly contradicting it without rebuke from the group? Gilbert suggests that members must make clear that they are speaking *in propria persona* in these cases, but we could easily imagine personal or professional consequences that would discourage individuals from voicing their dissent. For example, if a postdoc did this, the PI might not write them a strong letter of recommendation. On the other hand, if this were allowed the member's open contradiction could effectively undermine the group's position, and so their shared goal. Imagine, for example, an IPCC scientist speaking *in propria persona* that even though they believe in the IPCC's overall mission they think the consensus statement put out by the group is not epistemically important because it masks researchers' fruitful disagreement. Even if this is a tenable position, such a statement would be fodder for the climate skeptic. As Gilbert put it: "[…] it may make sense to worry about the effect her pronouncements could have on the community as a whole: perhaps its flourishing depends on its believing that [position]. Perhaps the [position] is fundamental to many of its projects and throwing it into doubt would lead to a period of crisis and even disintegration" (2000, 44).

[34] See Solomon (2006) and Tollefson (2006) for discussions of Janis's suggestions to mitigate groupthink. Also see Chapter 5 for a discussion of these strategies.

III A Purpose-Oriented Approach

So far I have presented two broad categories of consensus – aggregative and collaborative consensus – with models from the philosophical literature illustrating each. I also highlighted some strengths and weaknesses for each. Sometimes philosophers have framed aggregative and collaborative consensus as being in conflict (see Tollefsen 2006; Solomon 2001, 2006), and I want to suggest an alternative framing of their relation to one another. Conceptually, these models are making sense of different types of consensus in scientific practice as exhibited by the sorts of cases discussed in the literature (Beatty 2006; Fagan 2012; Staley 2007; Solomon 2006; Tollefsen 2006). What this suggests to me is that there simply are different types of consensus operating in scientific practice, especially since aggregative and collaborative consensus get used in quite different practical contexts. Therefore, the relationship between aggregative and collaborative consensus is not one of conflict and selection, but rather one about what sorts of purposes they are best suited for. Aggregative consensus, such as Solomon's normatively appropriate consensus model, can be useful for gauging the status of current theories in a research community writ-large. Some examples Solomon discusses include debates about the status of plate tectonics theory, the "central dogma" in molecular biology, and cancer research. Notice that all of these examples are about the status of a given theory in a specific discipline. Moreover, the similarity of expertise of members in these disciplines is crucial to providing insight on the status of these debates. This sort of model, to use List and Pettit's term, "pools information" by aggregating the judgments of members of a community, who are assumed to be competent about the issue under debate, to determine the current expert consensus on it. There are venues in contemporary scientific practice that do this, such as review papers that aim to depict the current "state of play" on a particular issue of interest for a scientific community, such as the *Annual Reviews* series (Beatty and Moore 2010, 213).[35]

On the other hand, collaborative consensus models such as Gilbert's joint commitment model or Tuomela's positional group beliefs, can be useful in contexts where groups organize themselves to meet explicit goals and promote interaction. Scientific bodies that make public policy recommendations are one example. They aim to provide concrete information on practical matters, such as health policies and climate strategies. The focus is less on "pooling information" among members, but on *working together* to meet shared aims. Furthermore, group members can have different expertise, and they can rely on each other's specialized knowledge. Their consensus position involves deliberation, negotiation, and reliance on others' expertise to realize joint goals. Collaborative consensus, unlike aggregative consensus, to use Beatty's (2017) words, sometimes "does not add up." The practical contexts in which aggregative and collaborative consensus occur in science are, in sum, different and they serve distinct purposes.[36]

[35] See the *Annual Reviews* website: https://www.annualreviews.org

[36] It should also be noted that within each category of aggregative and collaborative consensus particular models can also be better suited for certain purposes over others. For example, if a collaborative group

In addition, collaborative consensus models should not be assessed by assuming the methods employed by aggregative consensus models are most plausible, and vice versa. As we will see in Chapter 3, science skeptics confuse public discourse on climate change by implicitly invoking the techniques of aggregative consensus as appropriate normative standards for assessing collaborative consensus. This messaging then unfortunately makes its way to opportunistic public figures who borrow it to stall discussion and concrete action. Aggregative and collaborative consensus face their own unique challenges, and their remedies involve quite different resources. For example, if there is concern about unethical influence or manipulation during collaboration among members with unequal authority (e.g., between a PI and a graduate student), then an external committee could be created for group members to anonymously voice their concerns. Attention should be paid to the aims of the particular group, the context(s) from which they operate, and the sort of interaction(s) they partake to form consensus. It's not obvious that implementing a voting procedure for this sort of collaboration would resolve the issue. Getting clear on the difference between aggregative and collaborative consensus, as well as their strengths and weaknesses, allows for better analysis and evaluation of cases in practice.

That said, the techniques of aggregative and collaborative consensus models can also sometimes be jointly employed to establish a group's consensus position. John Beatty and Alfred Moore's (2010) "deliberative acceptance" model invokes (a) a collaborative component by promoting group deliberation with an emphasis on promoting the value of a persistent minority, and (b) an aggregative component by promoting a majority voting procedure to finalize the group position.[37] They define the account as follows:

> 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 stand as the position of the group. (2010, 209)

The authors emphasize the group's need for a persistent minority so that the "majority position is tested" (2010, 203). They argue that unanimity is problematic for there is a "temptation to submit to the conceit that one is infallible or that one's co-deliberators are vastly inferior" (2010, p. 204). Critical dissent towards the majority position can help deter such temptations. The inclusion of a persistent minority also presupposes "deliberative equality": objections from dissenters are adequately addressed because their points are worthy of consideration. In addition, majority voting is not understood as the greatest number ultimately winning out over other positions since this could be attributed to conformism like peer influence. It requires, instead, that members – and this includes the persistent minority – *agree* to

works from the assumption that all group members are equal, then Gilbert's model would be practically more useful than Tuomela's model.

[37] Note that the authors take Gilbert's model as their starting point. They modify Gilbert's model to incorporate an aggregative component – that is, a voting aspect requiring a majority. The authors write:

> Joint acceptance may not even involve a vote. Deliberative acceptance, on the other hand, does involve a vote, which is appropriate for deliberators who count each other as equals. To be sure, there is more to deliberative acceptance than just a vote. But, in this case, the count of votes matters. It is epistemically significant that the position jointly accepted was favored by a majority of deliberative equals. (2010, 210)

let *p* stand as the group's view. The majority vote is based on the quality of deliberation that took place beforehand. More specifically, the persistent minority felt that their concerns were fairly acknowledged and given sufficient treatment. However, Beatty and Moore point out that "deliberative agreement still extracts from the minority a significant concession, the acknowledgement that they were heard, but had not proven persuasive" (2010, 209).[38] But even though Beatty and Moore's account employs techniques from aggregative and collaborative models, the sorts of cases they have in mind are still primarily in the purview of collaboration since the groups they consider interact, through deliberation, to hash out a group position. The key difference is that a majority voting procedure finalizes the group position.

To summarize, aggregative and collaborative consensus need not be in conflict, and sometimes techniques from models of both categories can be used together. Philosophers are right to distinguish between aggregative and more collaborative types of consensus formation, but their being different does not amount to them being in conflict. They can sometimes overlap and even be mutually supportive.

Conclusion

The remainder of this dissertation will focus on offering a thorough analysis of collaborative consensus, including those that sometimes employ aggregative techniques in their decision-making procedures, with the aim of proposing normative conditions or guidelines to promote well-functioning collaborations that produce group positions. This project will be fruitful for several reasons. One reason is the centrality of consensus emerging from collaborations in current scientific practice such as jointly authored papers and public statements declaring a sub-community's or an entire community's expert consensus position on some issue such as climate change or vaccines. How should we understand these types of joint or collective endeavors in which members of a group share a group position? Finally, when and why should other communities trust collaboratively formed expert consensus? The public discourse about climate change, vaccines, or whether evolutionary theory and creationism should be taught side-by-side, often include discussion about whether there is dissent among experts. Science skeptics often appeal to disagreement among experts to promote doubt. If the consensus on some of these issues was established collaboratively, and not through a tallying of members' beliefs, when should people and their communities trust them?

In the next Chapter, I focus on Gilbert's joint commitment model, and some of the criticisms raised against it as a model for capturing collaborative consensus, given its wide-influence in the literature. In particular, I draw attention to Melinda Fagan's (2011; 2012) set of criticisms in two recent papers.

[38] Beatty and Moore write: "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 agree to let the position stand as the group's" (2010, 209).

Bibliography

Andersen, Hanne, and Susan Wagenknecht. 2013. “Epistemic Dependence in Interdisciplinary Groups.” *Synthese* 190: 1881–98.

Beatty, John. 2006. “Masking Disagreement Among Experts.” *Episteme* 3 (1–2): 16.

———. 2017. “Consensus: Sometimes It Doesn’t Add Up.” In *Landscapes of Collectivity in the Life Sciences*, edited by Snait Gissus, Ehud Lamm, and Ayelet Shavit, 179–89. Cambridge, MA: MIT Press.

Beatty, John, and Alfred Moore. 2010. “Should We Aim for Consensus?” *Episteme* 7 (3): 198–214.

Beaubien, Jason. 2020. “WHO Halts Hydroxychloroquine Trial Over Safety Concerns.” *NPR*, May 25, 2020, sec. Coronavirus Live Updates. https://www.npr.org/sections/coronavirus-live-updates/2020/05/25/861913688/who-halts-hydroxychloroquine-trial-over-safety-concerns.

Bratman, Michael. 1999. *The Faces of Intention*. Cambridge, MA: Cambridge University Press.

Collins, Harry, and Robert Evans. 2002. “The Third Wave of Science Studies: Studies of Expertise and Experience.” *Social Studies of Science* 32 (2): 235–96.

Fagan, Melinda Bonnie. 2011. “Is There Collective Scientific Knowledge? Arguments from Explanation.” *The Philosophical Quarterly* 61 (243): 247–269.

———. 2012. “Collective Scientific Knowledge: Collective Scientific Knowledge.” *Philosophy Compass* 7 (12): 821–31. https://doi.org/10.1111/j.1747-9991.2012.00528.x.

Fuller, Steve. 1986. “The Elusiveness of Consensus.”

Gilbert, Margaret. 1996. *Living Together*. Landham, MD: Rowman & Littlefield Publishers.

———. 2000. “Collective Belief and Scientific Change.” In *Sociality and Responsibility*, 37–49. Landham, MD: Rowman & Littlefield Publishers.

Janis, Irving. 1972. *Victims of Groupthink*. Boston: Houghton Mifflin.

Kitcher, Philip. 2001. *Science, Truth, and Democracy*. Oxford: Oxford University Press.

List, Christian, and Philip Pettit. 2002. “Aggregating Sets of Judgments: An Impossibility Result.” *Economics and Philosophy* 18: 89–110.

———. 2011. *Group Agency*. New York: Oxford University Press.

NIH. 2020. “Hydroxychloroquine Plus Azithromycin.” NIH COVID-19 Treatment Guidelines. July 17, 2020. https://www.covid19treatmentguidelines.nih.gov/antiviral-therapy/hydroxychloroquine-plus-azithromycin/.

Oprysko, Caitlin. 2020. “Trump Says He’s Taking Hydroxychloroquine, Despite Scientists’ Concerns.” *Politico*, May 18, 2020. https://www.politico.com/news/2020/05/18/trump-says-hes-taking-unproven-anti-malarial-drug-265546.

Oreskes, Naomi, and Erik M. Conway. 2010. *Merchants of Doubt*. New York: Bloomsbury.

Richardson, Alan. 2008. “Solomon’s Science without Conscience, or, On the Coherence of Epistemic Newtonianism.” *Perspectives on Science* 16 (3): 246–52.

Rolin, Kristina. 2008. “Science as Collective Knowledge.” *Cognitive Systems Research*, no. 9.

Schmitt, Frederick F. 1994. “The Justification of Group Beliefs.” In *Socializing Epistemology: The Social Dimensions of Knowledge*, edited by Frederick F. Schmitt. Maryland, USA: Rowman & Littlefield Publishers.

Schwartz, Matthew S., Susan Davis, Scott Detrow, and Juana Summers. 2020. “Democratic Candidates Step Up Criticism of Trump’s Handling of Coronavirus.” *NPR*, February 29, 2020, sec. The Coronavirus Crisis. https://www.npr.org/2020/02/29/810713730/democratic-candidates-step-up-criticism-of-trumps-handling-of-coronavirus.

Solomon, Miriam. 2001. *Social Empiricism*. Cambridge, MA: MIT Press.

———. 2006. “Groupthink versus the Wisdom of Crowds: The Social Epistemology of Deliberation and Dissent.” *The Southern Journal of Philosophy*, no. 44 Supplement: 28–42.

Staley, Kent. 2007. “Evidential Collaborations: Epistemic and Pragmatic Considerations in ‘Group Belief.’” *Social Epistemology* 21 (3): 15.

Surowiecki, James. 2004. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. New York: Random House.

Tuomela, Raimo. 1992. “Group Beliefs.” *Synthese* 91 (3): 285–318.

Urfalino, Phillipe. 2006. “Apparent Consensus and Voting: Modes of Collective Decision-Making.” https:// www.academia.edu/4277481/Apparent_Consensus_and_Vote.
Wray, K. Brad. 2001. “Collective Belief and Acceptance.” *Synthese* 129: 319–33.
Wylie, Alison. 2006. “Socially Naturalized Norms of Epistemic Rationality: Aggregation and Deliberation.” *The Southern Journal of Philosophy*, no. 44 Supplement: 43–48.

2 Resources for Analyzing Collaborative Consensus: Evaluating Gilbert's Joint Commitment Model

In the previous chapter, I sketched two broad categories of consensus relevant for scientific practice – aggregative and collaborative consensus. In this chapter, I want to focus discussion on one of the collaborative consensus models surveyed: Gilbert's (1996) joint commitment model. Gilbert's model is arguably the most widely discussed model in the philosophy of science and social epistemology of science literature. Some proponents consider it to be an exemplary model with conceptual resources that help us understand consensus of the collaborative sort (see Beatty 2006; Staley 2007; Weatherall and Gilbert 2016; Wray 2018). Kent Staley (2007), for example, uses the model to understand the Collider Detector at Fermilab (CDR) case, a large-scale collaboration in experimental physics that published a jointly authored paper with 450 co-authors in which they assert a group consensus position on supporting evidence for the top quark. However, there are also some philosophers who consider Gilbert's model to be imperfect or potentially problematic (Fagan 2011, 2012; Thagard 2010). Paul Thagard, for example, considers joint commitments to be a "rather rare" phenomenon in science and they are, at best, "metaphorical pointers" for underlying mechanisms (e.g., neural, psychological, and social), which Gilbert's model does not cover. He is thus skeptical of the "overly realist and ontologically puzzling attribution of beliefs and other mental states [the joint commitment model posits] to scientific groups and organizations" (Thagard 2010, 280).[39] To-date though, Melinda Fagan (2011; 2012) has offered the most comprehensive argument against Gilbert's joint commitment model, which echoes some of Thagard's concerns. Fagan questions the way Gilbert's model makes sense of the relation or connection between a group's consensus position and the beliefs of its individual members. She proposes an alternative "relational account," which gives special attention to the community's relations of epistemic authority. My aim in this chapter is consider Gilbert's model in more detail by engaging with Fagan's critique against it. I focus on Fagan's critique for both its comprehensiveness and for the way in which it specifically covers the application of Gilbert's model to scientific groups that form consensus. Moreover, a discussion of Fagan's critique will make explicit concerns that I suspect other scholars engaged with

[39] It should be noted that most proponents of the model (including Gilbert) now standardly understand groups to be engaging in joint *acceptance* and not beliefs in the context of science. Thagard is also weary of Gilbert's way of explaining joint commitments through everyday concepts or conventions (e.g., readings groups, library committees). He writes: "I do not share the view of many analytic philosophers that the goal of philosophy is just to describe ordinary concepts […] I think that philosophy, like science, should be revisionary, looking to develop concepts with greater explanatory potential than everyday ones" (2010, 280).

Gilbert's work have. In my analysis, I will argue that a proponent of Gilbert's model can adequately respond to Fagan's critique. Furthermore, my interest in this dissertation is to consider Gilbert's model in the context of scientific collaborations. So even if Fagan's critique is accurate with respect to the sorts of cases she has in mind – scientific disciplines – I don't think it undermines attempts to apply Gilbert's model to deliberately coordinated groups where members knowingly taking part in joint projects together.

This chapter will be organized as follows. In section I, I start with a summary of the contents of Gilbert's joint commitment model. In section II, I lay out Fagan's critique of Gilbert's model. In section III, I argue that a proponent of Gilbert's model can address Fagan's critique. And in section IV, I articulate why it would be valuable to keep the various models that can account for collaboratively produced consensus available as resources since scientific collaborations can vary widely in terms of organizational structure and social dynamics. The different features highlighted by these models can be useful when considering different cases.

I Gilbert's Joint Commitment Model

On the joint commitment model, a group position is formed when members of a group agree to let a position stand as the group's view. This model emphasizes acceptance where one can accept a position for a variety of reasons (e.g., through negotiation or for the sake of a broader goal). This is different than emphasizing the individual's belief about the truth or falsity of the group's position (see Wray 2001). Group members can thus have personal beliefs that differ from the group's position that they accept. Here is Gilbert's formal definition of the joint commitment model:

> (1) A group G [accepts] *p* if and only if members of G jointly accept that *p*.
>
> (2) Members of a group G jointly accept that *p* if and only if it is common knowledge in G that the individual members of G have openly expressed a conditional commitment jointly to accept that *p* together with the other members of G. (1996, 204-205)

Let me illustrate with Gilbert's discussion of a poetry reading group. The group has a joint commitment on a specific interpretation of a poem if and only if the group jointly accepts that interpretation. They "jointly accept" that interpretation if everyone in the group knows that together they all agree to stipulate that interpretation. The joint commitment is common knowledge. Furthermore, members can agree to let the position stand as the group's view for different reasons. Some members may personally disagree with the interpretation, but they accept it to, say, help deter irresolvable arguments about which interpretation is correct. Members' acceptance of one interpretation of the poem also entails normative obligations. They are, for instance, expected to take that interpretation as a starting point for group discussion. They thus cannot constantly bring up other competing interpretations, for members would respond with: "But we already agreed on an interpretation."

A distinctive feature of the joint commitment model is the way in which group positions are characterized as "irreducible" from the personal beliefs of individual group members.[40] Gilbert writes: "*it is not a necessary condition of a group's [acceptance] that p that most members of the group believe that p*. Indeed, […] it seems that *it is not necessary that any members of the group personally believe that p*" (1996, 201; original emphasis). If we consider the poetry reading group, their decision to let one interpretation stand as the group's position does not require that some or most group members personally believe it. The model is not summative. Indeed, the group can jointly accept the interpretation even if most or even all members do not personally believe it.[41] Gilbert points out that group members can let a position stand as the group's view for a variety of reasons, and this will depend on group dynamics. For example, perhaps the majority of members of the poetry reading group are indifferent about which interpretation to choose. They simply do not have an opinion about the matter. One person, however, does have a very strong personal belief about the best interpretation, so everyone goes along with it. In another scenario, if some group members are more vocal about advocating their personal beliefs, then other members might be reluctant to speak up and express their preference. All the other members may thus be willing to jointly accept the interpretation that opinionated group members prefer. Finally, in a different scenario, perhaps group members simply disagree about which interpretation of the poem to work with, so they compromise on an interpretation that no member advocated for but are all willing to accept.

Group positions are also an important phenomenon in scientific practice. Thomas Kuhn (1962/1996), for example, emphasized that during periods of normal science scientific communities (i.e., scientific disciplines) are organized primarily by way of consensus about fundamentals of a research paradigm. More recently, some philosophers of science have focused on consensus in scientific collaborations. Some examples include research teams that reach consensus on a general thesis or position of an article intended for publication (Andersen 2010; Staley 2007), or reaching a consensus position to help guide public policy (Beatty 2006). In these contexts, the conceptual resources of Gilbert's model can be a useful starting point to understand how consensus positions relate to individual group members and what sorts of obligations are entailed when members take part in joint endeavors.

II Fagan's Critique

But even if the joint commitment model can help make sense of group positions in scientific practice, is it the best or most preferable model for this purpose? Fagan (2011; 2012) contrasts the joint commitment model with relational models that foreground relations of epistemic authority and argues

[40] As we will see this feature is important to Fagan's critique.
[41] In Gilbert's words: "That *A* and *B* jointly accept that *p* neither entails nor is entailed by the individually accepting that *p*. It is important to stress this" (1996, 205).

that relational models can do this better. She does this by considering an example: molecular biology's Central Dogma. Fagan writes:

> It is common to suppose, for example, that there is consensus in the molecular biology community that biological information flows only from DNA to RNA to protein (i.e., the Central Dogma), just in case most molecular biologists believe this. […] [However,] [m]ost molecular biologists today do not believe the Central Dogma, in a strict sense. Yet the idea that biological information follows in a linear track from DNA to RNA to protein is still prevalent, though few molecular biologists would endorse it if pressed." (2012, 825-826)

Roughly, most molecular biologists do not endorse the Central Dogma, but it is assumed that there is a consensus in the community about it. Furthermore, this view is standardly taught to students. This case is one where the group's consensus position (i.e., the Central Dogma) does not align with the beliefs of its individual members (i.e., that the Central Dogma is technically false). So how might a proponent of the joint commitment model make sense of this example? Recall that a distinguishing feature of the joint commitment model, in Fagan's words, is that "some scientific knowledge is irreducibly had by groups" (2012, 822). That is, the group position need not match what all or most individual members personally believe – they can be different. A proponent of the joint commitment model might thus explain that the Central Dogma is a consensus position that is "irreducibly had" by the molecular biology community even though most molecular biologists do not believe it.

However, Fagan contrasts the explanation offered by a proponent of the joint commitment model with, broadly construed, relational models (see Tuomela 1992; Corlett 1996; Fagan 2011). These models emphasize relations of epistemic authority such as trust and authority among members, as well as the connection between members' beliefs.[42] They also do not emphasize the irreducibility of the group's position with members' personal beliefs. On relational models, the group's consensus position is neither an aggregation of individual members' beliefs, nor a joint commitment according to Gilbert's articulation. Rather, it is a system "of members' beliefs and social epistemic relations [...]" (2012, 826). Fagan writes:

> The relational view provides a plausible account of the Central Dogma: its persistence in molecular biology today is fully determined by molecular biologists' beliefs, together with relations of epistemic authority that structure the molecular biology community. In this case teaching practices are important. All or most students of molecular biology are taught, by experts whose authority they accept, that the Central Dogma is (basically) correct. So the idea persists, though students who go on to become practicing molecular biologists later learn that it is, strictly speaking, false. (2012, 827)

[42] Note that Fagan's articulation of relational models is fairly preliminary, and the quotes offered here are most of it. At the time of our last correspondence, she was still working on and thinking about offering a more robust articulation of this model.

So the Central Dogma persists for educational purposes, and not because the molecular biology community, as a group, has a belief that is irreducible from the personal beliefs of its individual members. Fagan emphasizes the importance of teaching the Central Dogma by authoritative experts. Though the experts do not endorse the Central Dogma, for pedagogical purposes they teach it to all molecular biology students. Students who further study the subject eventually learn that the Central Dogma is false.

The joint commitment model and relational models can thus both explain why the Central Dogma persists even though most molecular biologists believe it to be false. But Fagan argues relational models offer a better explanation for the Central Dogma's persistence. A crucial reason why is that on Gilbert's model joint commitments are *irreducible* in the sense that they are not connected to the beliefs of group members. Fagan writes: "[t]he very features that render joint commitment irreducible - it's simplicity, unity, and independence from a complex of individuals' epistemic attitudes - make it's connection to those attitudes mysterious" (2012, 828). Fagan goes on to say:

> This is not to say that [the joint commitment model] provides no explanation whatsoever of social phenomena. But it cannot provide a better explanation than an alternative that omits the concept of joint commitment. Any social phenomenon that can be explained in terms of plural subjects of joint commitment can be explained more simply without irreducible joint commitment, in terms of the network of obligations and entitlements that structures relations of members' beliefs to one another - that is, in terms of relational belief. (2012, 828)

Relational models, according to her, are simpler. They focus on the interactions among community members with respect to their positional authority. The Central Dogma's persistence is explained by looking at the molecular biology community's relations of epistemic authority and their teaching practices. We do not need to appeal to the phenomenon of joint commitments to make sense of this example. In this sense, Fagan argues, relational models are preferable because they can explain the Central Dogma's persistence without positing an entity above and beyond the beliefs of the molecular biology community's members.

In the next section, I offer a response to Fagan's critique. I argue that a proponent of the joint commitment model can adequately respond.

III The Proponent's Response

A central point of focus in Fagan's critique of the joint commitment model pertains to the irreducibility of the group position in relation to group members' personal beliefs. However, I think a proponent of joint commitment who is interested in applying the model to scientific practice can ameliorate Fagan's concerns. When we talk about group positions formed through joint commitments, because they are irreducible from the personal beliefs of individual group members, it may seem like the group is being improperly treated, metaphysically speaking, "as a thing" independent from anyone

(Gilbert 2000, 38). However, the proponent might point out that while the irreducibility feature does suggest that the group position need not be connected to group members' personal beliefs, this does not mean the group position is therefore completely disconnected from group members – they are connected in some other sense. If we look a little more closely at Gilbert's writing, this last point becomes clearer. She writes:

> I do not imagine a community's existence and thought to be so far removed from the individual members of the community that their states of mind are not relevant to it. […] [Joint commitments] are a matter of how it is with the individual members of the group in question. What it is not necessarily a matter of, however, is the corresponding beliefs of those members. That is, it is not the case that the group [accepts] that *p* if and only if all or most members believe that *p*. [...] To say that, however, is not to say that a group's [position] is not a matter of how it is with the members. In the clearest case, *something* will be true of all members" (2000, 39; original emphasis).

From this quote, we see that Gilbert does intend there to be an important relation between the group position and the individual group members. This is achieved "when, roughly, each of the parties has expressed his or her personal willingness to be party to [a joint commitment] in conditions of common knowledge" (2000, 40). Moreover, once group members have formed a joint commitment, they are "bound together" in a significant way. This is apparent when we consider individual group members' obligations associated with the joint commitment. If we consider again the poetry reading group, once the group has settled on an interpretation of a poem, all participants – in their capacity as group members – are expected to work with that interpretation. If one member continues to object even though they agreed to take on the group's chosen interpretation, they would understandably be met with rebuke from other group members. Thus, on Gilbert's model, the connection or relation between the group position and individual group members pertains to each members' willingness to agree to let some position stand as the group's and it being common knowledge. Furthermore, when each individual is acting in their capacity as a group member, even if they personally do not believe the group position, they understand their obligation to uphold it. So while it is true that joint commitments can be independent from individual members' personal beliefs, it does not mean that there is no relation between the joint commitment and the individual members.

It is worth mentioning that people take on joint commitments and act on them even if they personally believe otherwise in all aspects of social life. It is an everyday social phenomenon that should be familiar to most of us. In scientific contexts, researchers make compromises with their team members that they may not personally believe to put out joint products, such as a co-authored paper or a conference presentation. In academic departments, members can jointly choose to offer a tenure-track position to someone no one in the department had initially recommended. This can happen, for example, in cases where department members are split on the other candidates, so they choose the least controversial person. This is just to say that joint commitments should be rather familiar to most us.

Given this, I think Fagan's critique that the joint commitment model is "mysterious" misses the mark. A proponent can point out that the connection between the group position and individual group members is located in members' willingness to let some position stand as the group's.

Perhaps Fagan will point out that even if joint commitments are not mysterious, they still posit one more entity compared to relational models. Relational models offer a simpler explanation. However, though simplicity can be an epistemic virtue (see Kuhn 1977) in some explanatory contexts, it is not clear why it should be used to compare and assess the models here, especially since the joint commitment model is not mysterious. If neither model makes problematic assumptions and they can both adequately explain the phenomenon under consideration, then it is unclear why the simpler theory is better. The burden of proof I think rests with those making appeals to simplicity.

Perhaps Fagan might argue that relational models, by looking at relations of epistemic authority, actually do better explaining the persistence of the Central Dogma in the molecular biology community by highlighting that it is intended for educational and pedagogical purposes. However, even here I think a proponent of the joint commitment model can challenge Fagan's analysis by pointing out problems with the example. Recall that on Fagan's analysis, a supporter of the joint commitment model makes sense of the example by explaining that the molecular biology community's consensus position on the Central Dogma persists even though individual practitioners believe it is false because the group's position is irreducible to the beliefs of individual members. But even though this irreducibility feature is central to the model, there's more to it. A group that partakes in joint commitment is importantly doing it *together* and there are obligations entailed when such a commitment is made. Some story would need to be told about how the molecular biology community (or perhaps some authoritative subset of the community) jointly agreed that teaching the Central Dogma is important for their curriculum and pedagogical goals for training future practitioners. Members responsible for teaching would also need to know that it is common knowledge among the group that they have agreed to let the position stand as the group's view. And we would need to know whether they have an obligation to uphold the consensus position. Assessing joint commitments cannot rely solely on there being an irreducible group view that is separate from the beliefs of individual members. It would also need to be shown that members of the molecular biology community have jointly agreed to teach the Central Dogma to students, even though practitioners do not believe it. But this has not been shown. In this sense, it seems Fagan's assessment of the joint commitment model with this example is incomplete.

Relatedly, a proponent of the joint commitment model might reject Fagan's argument if they take seriously K. Brad Wray's (2007; 2018)[43] claim that only research teams can have joint commitments, while research disciplines, sub-fields, and the scientific community more generally cannot.[44] Wray argues that research teams possess what Emile Durkheim calls "organic solidarity." These are groups where

[43] Also see Wray's (2018, 123) response to Fagan's critique. He argues that her view is actually consistent with Gilbert's model.
[44] Wray is often considered a proponent of Gilbert's joint commitment model.

members share a functional dependence in which they rely on each other to perform their assigned function. A group of factory workers, for example, possess organic solidarity because members are functionally dependent on each other to perform assigned functions. The group cannot complete their "end product" without everyone performing their function in the division of labor (2007, 342). Similarly, research teams – especially interdisciplinary ones – possess organic solidarity: members perform their assigned function while deferring to others on matters they do not know – e.g., the engineer will defer to the neuroscientist on matters relating to the nature of the brain, and the neuroscientist will defer to the engineer on the design and construction of neurotechnologies such as brain-computer interfaces. Because of this functional dependence among group members, research teams have the "capacity to adopt a view that is *irreducibly the group's view*" (2007, 343; original emphasis). Research fields, sub-fields – e.g., the community of astrophysicists and the community of endocrinologists - and the scientific community in general, on the other hand, possess what Durkheim calls "mechanical solidarity," not organic solidarity. These groups can have "similar thoughts and attitudes" but as individuals (2007, 341). They are not organized as a "unified whole" in the same way that research teams are, nor are the members functionally dependent on each another in general. If a proponent of the joint commitment model takes this line of reasoning seriously, then they may also object to Fagan's use of the Central Dogma as an example because the molecular biology community is a research field that can only have mechanical solidarity, but not organic solidarity. They therefore cannot make joint commitments.[45] Note that this would be an additional stipulation to Gilbert's model, and it would constrain what sorts of groups it applies to in science.

In sum, there is space for a proponent of the joint commitment model to remain skeptical of Fagan's critique. They may emphasize that the connection between the group position and the beliefs of individual members is not "mysterious" if one pays attention to the actual formation of joint commitments and the obligations that group members are bound to when they make joint commitment. In addition, they may question the completeness of Fagan's discussion where she applies the joint commitment model to the Central Dogma. In that example she only focuses on there being a group position that is irreducible to the personal beliefs of molecular biologists. But as I discussed, there is more to Gilbert's model than just the irreducibility feature. Fagan would need a full application of the joint commitment model to the example to make the case that relational models are better. In addition, the proponent may alternatively make use of Wray's distinction and point out that the molecular biology is a research field that possess mechanical solidarity for which joint commitments cannot be formed.

However, even if a proponent of Gilbert's model can respond to Fagan's critique, my view is that we should not focus our attention on using only the joint commitment model to make sense of

[45] Note that if, hypothetically speaking, some education task force was created for the molecular biology field to help standardize curriculum and they decide to implement teaching the Central Dogma for every intro class, then this response using Wray's distinction may not work.

collaboratively produced consensus in science. In the next section, I advocate to keep a plurality of collaborative consensus models available to assess actual cases.

IV Discussion

Even though to my mind a proponent of Gilbert's joint commitment model can adequately respond to Fagan's critique, this does not mean it is the only or most viable model with conceptual resources to make sense of collaborative consensus in science. Indeed, not all groups that produce consensus of the collaborative sort are identical, especially with regard to their organizational structure and social dynamics (see Cho 2011; Wray 2018). The collaborative consensus models discussed previously highlight important features that can vary across groups (see Chapter 1). Gilbert's model draws attention to how a group can agree to let a position stand as the group's view in a way that entails a normative obligation for members to promote and uphold that position as group members. On the other hand, Tuomela's positional group belief model draws attention to different social roles group members can occupy, and how some members can possess more authority than others.[46] And if we take the features of both Gilbert's and Tuomela's models into consideration, then the normative obligations for group members will differ depending on their social role in the group. Beatty and Moore's deliberative acceptance model draws attention to the importance of a persistent minority, especially in deliberative contexts, to help ensure that the group has considered a variety of perspectives. Pragmatically speaking, it seems valuable to keep these features on the table as resources to help understand collaborative consensus in practice. Some groups will possess all or some combination of these features, while others may possess only one. And even though I disagree with Fagan's critique of Gilbert's model, her emphasis, in agreement with Tuomela's model, on authority structures is important to keep in mind. My general position is that these consensus models are all potentially useful to understand collaborative consensus as a social phenomenon in scientific practice. But any analysis of collaborative consensus will depend on the group that is intentionally partaking in a joint activity together in scientific practice. Let me now briefly consider two examples (one of which I already discussed last chapter) of groups that have produced collaborative consensus positions as a result of their joint endeavour. The examples will show why it would be an advantage to keep the various collaborative consensus models we have considered available for consideration.

First, consider the interdisciplinary group the Neurotechnology Ethics Task Force (Yuste and Sara Goering et al, 2017) which recently published in *Nature* a co-authored paper about the ethics of neurotechnologies, which involve brain-computer interfaces (BCIs) and machine learning. Briefly, BCIs

[46] It is important to note that recently Tuomela (2013) views his model and being consistent and perhaps even an extension of Gilbert's model. This was not the case when he initially introduced his model in 1992 (see Tuomela 1992).

are neurotechnologies that have the capacity to record and/or stimulate neural activity from the brain, and then translate the information to an external device such as a computer curser or a prosthetic limb. This interdisciplinary group generating this paper has 25 co-authors working in neuroscience, philosophy, engineering, medicine, government agencies, tech industry, and the humanities more generally.[47] The group cautions against the dangers neurotechnologies like BCIs pose to society writ-large, especially regarding issues related to privacy and consent, agency and identity, augmentation, and bias. Organizationally, the group has two lead authors (Yuste, a neuroscientist and Goering, a philosopher) who are responsible for overseeing the entire paper and combining the sections written by different working groups. Everyone else is divided into four working groups, and each group is responsible for writing their assigned section of the paper. There is thus a division of labor for the writing. There is also a division of cognitive labor with respect to expertise. The philosophy and humanities scholars enlist their expertise on ethics and social justice scholarship. The scientists, researchers, and tech industry members enlist their expertise about the most recent developments and capabilities of neurotechnologies taking place in their respective fields. Government officials enlist their expertise on the sorts of public policies that governments have or could potentially enact in their respective nation states. Once each working group finishes writing their assigned sections, they submit the content to the lead writers who then put everything together. The lead writers make executive editorial decisions to improve the paper's argumentative flow and provide a framing and conclusion. Collectively, the authors are motivated to publish a piece of writing that discusses the need for ethical guidelines for neurotechnologies given their impact on people's privacy, and their sense of self, as well as their relationships with others. They see their group's interdisciplinary collaboration as a strength because it shows that experts in a variety of different areas recognize the need for further discourse and public policies that protect end users, and society more generally, from the proliferation of neurotechnologies. Note that the fact that this is an interdisciplinary group means that members do not share the same expertise, and they defer to each other on subjects they do not know. More specifically, given the range of expertise, there will be parts of the paper that only some, perhaps few, group members will be able to evaluate. And yet, the consensus position in the published paper about which ethical principles and issues to prioritize for neurotechnologies nevertheless reflects the group's collaboratively produced view. Each member contributed to the writing, and they were each asked to provide consent to let the position in the final version of the paper stand as the group's position. Members were also given the option to opt out if they were not comfortable with the contents of the paper, and some did.

Now contrast the example above with Beatty's (2006) historical example that I discussed last chapter about the group of geneticists who reach consensus on an appropriate threshold for radiation exposure – 10 *roentgens* (see Chapter 1). Recall that there was substantial disagreement among the geneticists on how to make this calculation, even though they all shared the same disciplinary

[47] I should note that this example is one for which I was a co-author.

background. This historical example and the interdisciplinary co-authorship example just mentioned are both cases of collaborative consensus because the groups have explicit goals that require members' collective efforts to realize them. The geneticists are concerned about preserving scientific authority and protecting members of the public from the potential harms of radiation exposure. The interdisciplinary researchers are concerned about the potential harms of neurotechnologies on end users (given their rapid research expansion and development) and a lack of critical and ethical scrutiny. These consensus positions were produced as joint efforts. However, it is also important to take note of these groups' differences. The geneticists share the same expertise (even if they specialize on very specific topics in their field), while the interdisciplinary researchers do not. Moreover, the neurotechnology group is structured hierarchically with lead writers (who are also senior researchers) in charge while the geneticists are generally considered equals. I note these differences to illustrate that groups that produce collaborative consensus can have quite different social dynamics and organizational structures.[48]

In considering the two examples above, it seems an advantage to keep various collaborative consensus models on the table. As Beatty's analysis shows, Gilbert's model provides conceptual machinery to explain how the geneticists, despite their deep internal disagreements, manage to reach a consensus position on a permissible radiation dosage for policymakers. By contrast, the interdisciplinary co-authorship case requires drawing resources offered by both Gilbert's and Tuomela's models.[49] Gilbert's joint commitment model helps explain how the group of collaborators can share a position even though they do not possess the same expertise: by each jointly consenting to let the final version of the paper stand as the group's position. Tuomela's positional group belief model helps explain how in this case some group members have more authority than others – namely the lead authors. Group members grant the lead authors authority to make calls for the group as a whole, such as their efforts to fit the different sections written by each working group together to form a complete paper. The lead authors are also given executive authority to respond to reviewer comments on behalf of the group. So while I think Gilbert's joint commitment model can be extremely helpful for making sense of particular cases of collaborative consensus, I think it would be an intellectual loss to foreclose consideration of other models given the variation of cases from actual practice. Any analysis of consensus in scientific practice should be informed by examples from scientific practice, and not just philosophical debates.

[48] As I discuss in Chapters 4 and 5, this matters when we evaluate whether a collaboratively-produced consensus position is legitimate or trustworthy.

[49] I mention Tuomela's model here and not Fagan because she does not offer her own model but instead emphasizes a variety of models that pay attention to epistemic authority. One of the models she supports is Tuomela's.

Conclusion

In this chapter, I focused my discussion on Fagan's critique of Gilbert's joint commitment model, especially in its application to cases of scientific consensus. I argued that a proponent of Gilbert's model can adequately address Fagan's critique. But even so, this does not mean that Gilbert's model should always be preferable to other models, especially since different models emphasize slightly different features that are relevant to real cases. We should keep these differences on the table because the social dynamics and organizational structures of research groups are not uniform – they can vary widely.

On a broader note, my main goals and interests in this dissertation do not lie in resolving conceptual or theoretical quibbles between philosophers and social epistemologists of science. Rather, my goals and interests are geared towards making use of the philosophical contents discussed to contribute to public debates about scientific expertise, especially those concerning expert scientific consensus. In the next chapter, I offer an analysis of the debate on climate change using some of the ideas discussed so far.

Bibliography

Andersen, Hanne. 2010. “Joint Acceptance and Scientific Change: A Case Study.” *Episteme* 7 (3): 248–65. https://doi.org/10.3366/epi.2010.0206.

Beatty, John. 2006. “Masking Disagreement Among Experts.” *Episteme* 3 (1–2): 16.

Beatty, John, and Alfred Moore. 2010. “Should We Aim for Consensus?” *Episteme* 7 (3): 198–214.

Bratman, Michael. 1999. *The Faces of Intention*. Cambridge, MA: Cambridge University Press.

Cho, Adrian. 2011. “Particle Physicists’ New Extreme Teams.” *Science* 333 (6049): 1564–67.

Cortlett, J. Angelo. 1996. *Analyzing Social Knowledge*. Landham, MD: Rowman & Littlefield Publishers.

Fagan, Melinda Bonnie. 2011. “Is There Collective Scientific Knowledge? Arguments from Explanation.” *The Philosophical Quarterly* 61 (243): 247–269.

———. 2012. “Collective Scientific Knowledge: Collective Scientific Knowledge.” *Philosophy Compass* 7 (12): 821–31. https://doi.org/10.1111/j.1747-9991.2012.00528.x.

Gilbert, Margaret. 1987. “Modelling Collective Belief.” *Synthese* 73 (1): 185–204.

———. 1996. *Living Together*. Landham, MD: Rowman & Littlefield Publishers.

———. 2000. “Collective Belief and Scientific Change.” In *Sociality and Responsibility*, 37–49. Landham, MD: Rowman & Littlefield Publishers.

Kuhn, Thomas. 1977. “Objectivity, Value Judgments, and Theory.” In *The Essential Tension*, 320–39. Chicago: University of Chicago Press.

Staley, Kent. 2007. “Evidential Collaborations: Epistemic and Pragmatic Considerations in ‘Group Belief.’” *Social Epistemology* 21 (3): 15.

Thagard, Paul. 2010. “Explaining Economic Crises: Are There Collective Representations?” *Episteme* 7 (3): 266–83.

Tuomela, Raimo. 1992. “Group Beliefs.” *Synthese* 91 (3): 285–318.

———. 2013. *Social Ontology: Collective Intentionality and Group Agents*. Oxford: Oxford University Press.

Weatherall, James Owen, and Margaret Gilbert. 2016. “Collective Belief, Kuhn, and the String Theory Community.” In *The Epistemic Life of Groups*, edited by Michael S. Brady and Miranda Fricker, 191–217. Oxford: Oxford University Press.

Wray, K. Brad. 2001. “Collective Belief and Acceptance.” *Synthese* 129: 319–33.

———. 2007. “Who Has Scientific Knowledge.” *Social Epistemology* 21 (3): 337–47.

———. 2018. “The Impact of Collaboration on the Epistemic Cultures of Science.” In *Scientific Collaboration and Collective Knowledge*, edited by Thomas Boyer-Kassem, Conor Mayo-Wilson, and Michael Weisberg, 117–34. New York: Oxford University Press.

Yuste, Rafael, Sara Goering, and et al. 2017. “Four Ethical Priorities for Neurotechnologies and AI.” *Nature* 551: 159–63.

3 Considering Consensus in the Climate Change Controversy

In previous chapters, I mentioned that one of the ways in which science skeptics undermine the recommendations of science experts, especially to stall public policies they disagree with, is by pointing out any apparent lack of consensus among experts.[50] In this chapter, I focus on the very public and fraught debate about the status of anthropogenic climate change. The stakes are high because if we take anthropogenic climate change seriously then both national and international public policies would be required to curb emissions (e.g., decrease reliance on oil, coal, and fossil fuels, and implementing a carbon tax) while also promoting energy efficient options (e.g., solar energy, electric cars, reusable materials). Climate change deniers, however, highlight lack of agreement among climate scientists regarding anthropogenic climate change (see Oreskes and Conway 2010). A crucial, though implicit, assumption behind their strategy is the idea that expert opinion is trustworthy only if there is near or complete unanimous consensus. Let's call this the "unanimity criterion" for trustworthy scientific consensus, which, as we will see, is some skeptics' threshold for legitimate aggregative consensus (see Chapter 1). Groups such as the "Petition Project" purport to have signatures of "scientists," members of the same expert community, who reject evidence in favor of climate change.[51] As of June 2020, they claim to have 31,487 who have signed on, 9029 of which have PhDs. They write that "an overwhelming 'consensus' in favor of the hypothesis of human-caused global warming and consequent climatological damage is wrong. No such consensus or settled science exists. As indicated by the petition text and signatory list, a very large number of American scientists reject this hypothesis" (Petition Project).[52] Based on this reasoning, one should be skeptical of anthropogenic climate change since there is substantial dissent within the expert community. The skeptics' appeal to the unanimity criterion aims to sow doubt and confusion about anthropogenic climate change, especially in the public sphere. US public opinion polls between 1997 and 2007 show that 60% of the public believe there is substantial disagreement among climate scientists about the reality of anthropogenic climate change (Cook et al. 2013, 2; Nisbet and Meyer 2007). Given the centrality of expert 'scientific consensus' in this debate, we need to pay attention to how the skeptics use this notion to respond. Better understanding could also help public discourse in a politically charged climate with no shortage of alternative facts and explanations, as well as powerful corporations who lobby politicians in charge of crafting public policies to stall actions that would hurt their bottom line.

[50] 'Consensus' and 'agreement' will be used interchangeably.
[51] See "Global Warming Petition Project" here: http://www.petitionproject.org
[52] http://www.petitionproject.org/purpose_of_petition.php

A prominent and well-known defender of climate science research, Naomi Oreskes (2004a; 2004b), has very publicly responded to skeptics, especially their critiques of the Intergovernmental Panel on Climate Change (IPCC). The IPCC is a body of climate science experts from around the globe that publishes comprehensive assessment reports on the current status of climate change every 5-7 years for policymakers, as well as the broader public.[53] They are perhaps most well-known for their unequivocal consensus position that anthropogenic climate change is real. This consensus position though, as we will see, has received much attention and downplaying from skeptics. In defending the IPCC, Oreskes conducted a survey that corroborates the IPCC's consensus position. She looks at the stated positions in the abstracts of 928 peer-reviewed publications with respect to climate change and then aggregates them. She shows that none of the articles surveyed challenge anthropogenic climate change. Such a response however effectively accepts the unanimity criterion. Despite Oreskes's good intentions, I argue that an important factor has not been considered in her analysis. Many of the papers surveyed are jointly authored and so many of the stated positions come from groups that partake in collaboration – not just individual researchers. This creates complications for Oreskes's argument because it is unclear whether jointly authored positions can be aggregated, and if so, how. Overall, I question both the skeptics' and Oreskes's acceptance of the unanimity criterion and argue that contrary to their focus on aggregative consensus, the IPCC's consensus position is a collaboratively produced one. In this sense, I think it is a mistake to try to meet the skeptics' demand that near or complete unanimous consensus be met.

The aim of this chapter is to consider whether the unanimity criterion – the skeptics' threshold for aggregative consensus – is an appropriate measure to trust (or not trust) the IPCC's consensus position on anthropogenic climate change. I will put to use some of the distinctions and points of emphasis I have made in the first two chapters of this dissertation to make sense of the discourse and argumentative strategies of both skeptics and Oreskes. In particular, I use aggregative and collaborative consensus to help disentangle some of the implicit assumptions behind their arguments. In section I, I briefly introduce the IPCC and explain why we should understand the declarations in their assessment reports as constituting a consensus position. In section II, I provide an example of how climate change skeptics challenge the IPCC's consensus position. I show how they use the unanimity criterion as a measure to assess whether an expert group's consensus position is legitimate, and then undermine it by pointing to dissenting opinions to show a contradiction. In section III, I discuss Oreskes's response to the skeptics, which aims to show that the unanimity criterion can be met. In section IV, I show that many of the publications considered in Oreskes's survey are jointly authored, and I argue the positions asserted in them cannot be straightforwardly aggregated. This raises questions about Oreskes's overall response to the skeptics. In section V, I argue the IPCC's consensus position does not fit with aggregative consensus, and for this reason near or complete unanimity is not an applicable criterion.

[53] More on the IPCC in section I.

I The Intergovernmental Panel on Climate Change (IPCC)

One expert body that has received much critical attention from climate change skeptics is the Intergovernmental Panel on Climate Change, more commonly referred to as the IPCC. The IPCC was created in 1988 by the World Meteorological Organization (WMO) and the United Nations Environment Programme (UNEP). The group, comprised of many climate scientists globally, publishes a joint assessment report every 5-7 years on the current status of climate change (see Figure 1).

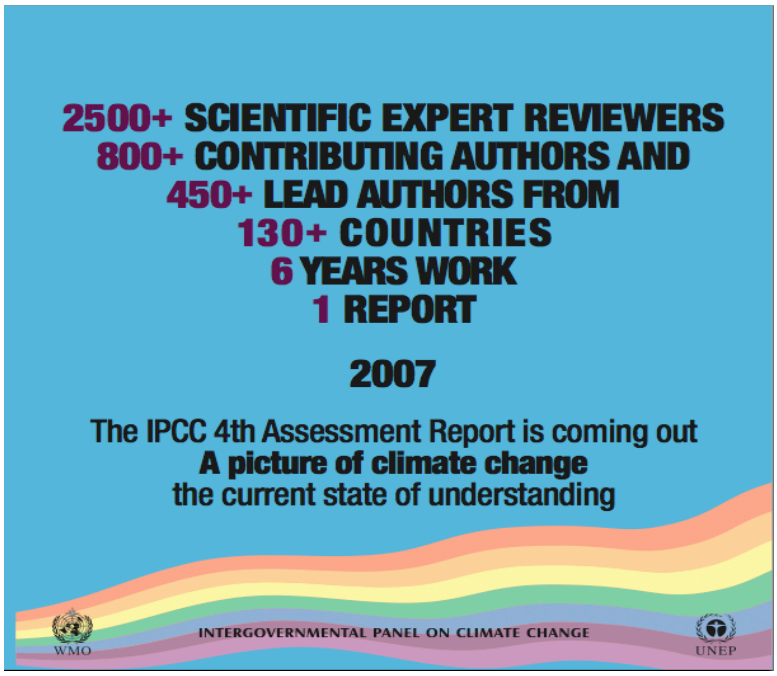

Figure 1: IPCC announcement on the 4th Assessment Report.

These assessment reports are intended to help with the development of climate policies. The group states that "The IPCC was created to provide policymakers with regular scientific assessments on climate change, its implications and potential future risks, as well as to put forward adaptation and mitigation options" (IPCC 2020b).[54] One defining feature of the IPCC group's report is its unequivocal position on the status of anthropogenic climate change. In their Fifth Assessment Report, Working Group I's Summary for Policy Makers states: "Warming of the climate system is unequivocal, and since the 1950s, many of the observed changes are unprecedented over decades to millennia. The atmosphere and ocean have warmed, the amounts of snow and ice have diminished, and sea level has risen" (IPCC 2013, 4). And later on in the Summary, they state: "Human influence on the climate system is clear. This is evident from the increasing greenhouse gas concentrations in the atmosphere, positive radiative forcing, observed warming, and understanding of the climate system" (2013, 15). These jointly authored assessment

[54] See: https://www.ipcc.ch

reports and the conclusions in their summary reflect the IPCC group's position as an expert body on the current status of climate change.

The IPCC's assessment reports have become associated with the view that there is a consensus among experts on the reality of anthropogenic climate change. John Beatty (2017) cites the science writer Cristine Russell's 2013 story, which reports that the Fourth Assessment Report "represent[s] a consensus from some 800 climate experts around the world who are commissioned to review the latest scientific evidence."[55] Interestingly though, the assessment reports themselves do not explicitly say 'consensus'. But there are a variety of statements and texts that strongly suggest that the declarations in the reports constitute a consensus position. Jean Goodwin's (2009) research draws attention to the IPCC's First Assessment Report in 1990, in particular the *Forward* to Working Group I's report, which focuses on the physical science basis. In that *Forward*, Chairman John Houghton writes:

> In preparation of the main Assessment most of the active scientists working in the field have been involved. One hundred and seventy scientists from 25 countries have contributed to it, either through participation in the twelve international workshops organised specially for the purpose or through written contributions. A further 200 scientists have been involved in the peer review of the draft report. Although, as in any developing scientific topic, there is a minority of opinions which we have not been able to accommodate, the peer review has helped to ensure a high degree of consensus among authors and reviewers regarding the results presented. Thus the Assessment is an authoritative statement of the views of the international scientific community at this time....I am confident that the Assessment and its Summary will provide the necessary firm scientific foundation for the forthcoming discussions and negotiations on the appropriate strategy for response and action regarding the issue of climate change. It is thus, I believe, a significant step forward in meeting what is potentially the greatest global environmental challenge facing mankind. (Houghton, Jenkins and Ephraums 1990, v-vi)[56]

Note Houghton's emphasis on there being a "high degree of consensus among authors and reviewers regarding the results presented." On Goodwin's rhetorical analysis, this consensus claim is an appeal to authority, especially intended for non-scientists. She also notes that this can be seen in some of the ways IPCC scientists have responded to skeptics in public reporting venues:

> Dr Jeff Jenkins, a Meteorological Office scientist and co-ordinator of the IPCC scientific group said yesterday: "This report is just not helpful. A number of things in it are misleading. The report we've issued is an international one, compiled by 170 scientists, the best ones we could find, and we would not expect their conclusions to be second-guessed by a report which is not as close to the scientific consensus as our is." (*The Independent* (London) 6/14/90)[57]

[55] *Grist*, https://archives.cjr.org/the_observatory/ipcc_coverage.php
[56] Also see Goodwin (2009) on this quote from the First Assessment Report's *Forward*.
[57] See Goodwin (2009, 5).

Notice again how in response to skeptics, Jenkins re-emphasizes the import of the IPCC's consensus position. Finally, and perhaps most importantly, the IPCC's procedures also tell us that they aim to reach consensus: "[in] taking decisions, drawing conclusions, and adopting reports, the IPCC Plenary and Working Groups shall use the best endeavours to reach consensus" (IPCC 2013c, 2). Given that IPCC leaders emphasize consensus in the assessment reports and consensus is explicitly stated as a goal in the group's procedures, these are good reasons to view the declarations in the assessment reports as reflecting the consensus position of the IPCC as an expert body on climate change. However, their highly publicized assessment reports and their unequivocal consensus position on anthropogenic climate change, also makes them a regular target for skeptics.

II An Argument from Skeptics

One of the ways in which skeptics try to undermine the IPCC's consensus position is to point out that the assessment reports are not trustworthy because there are scientists who disagree with it. One example is the *Heritage Foundation*, an influential conservative think tank. One of their authors writes the following:

> University of Virginia professor Fred Singer recently published an 800-page report titled, "Climate Change Reconsidered," which questions and debunks many of the IPCC conclusions and emphasizes that there is no scientific consensus on climate change. Richard Lindzen, professor of meteorology at the Massachusetts Institute of Technology, notes that the IPCC's models fail to take into account naturally occurring cycles such as El Niño, the Pacific decadal oscillation, or the Atlantic multidecadal oscillation. Other prominent scientists called political action "irresponsible and immoral" because of the lack of credible evidence. When the IPCC released its report in 2007, 400 climate experts disputed the findings; that number has since grown to more than 700 scientists, including several current and former IPCC scientists. (Loris 2010, 4)

Here, the author cites scientists, purportedly including IPCC scientists, who challenge the IPCC's consensus position on climate change. They appeal to these experts' dissent to motivate doubt about the IPCC's position. The argument is roughly as follows. The IPCC assessment reports cite evidence for anthropogenic climate change, which leads the group to a consensus position on the issue. However, the author argues, the consensus is not decisive because there are many hundreds of scientists who disagree with the evidence for the claim, including some who currently or formerly were part of the IPCC. A key assumption here is that one ought not believe expert consensus about anthropogenic climate change if there is in actuality substantial dissent about it.[58]

[58] Note that the experts cited include a physicist, climate scientists, and some current and former IPCC scientists. They offer evidence of dissent from scientists from a variety of scientific communities.

The sort of consensus the skeptic has in mind here should sound vaguely familiar. It is aggregative consensus. Every relevant expert gets something like a hypothetical vote on the issue, and then the votes are tallied. The total votes must meet some minimum threshold to count as acceptable (aggregative) consensus. The threshold is not specified but based on their argument above we know that substantial dissent is unacceptable. Later on in the article, the author writes: "If the vast majority of climatologists do not believe that the debate on climate change is over, politicians should not be pushing for greenhouse gas reduction policies that not only have significant economic costs, but will also deeply alter the business landscape of the United States" (Loris 2010, 7-8). 'The vast majority' here suggests a threshold of near or complete unanimity. But whatever the threshold, based on the number of expert dissenters cited in the article, matters do not look promising since hundreds (Loris 2010), and perhaps thousands if we take into consideration the signatures from the Petition Project earlier, disagree with the IPCC's consensus position. This line of argumentation is quite common in mainstream media that is skeptical of "big government" policies and of scientific research that could lead to broad public policies. It grants that legitimate expert scientific consensus is possible, but the existence of vast disagreement means no such consensus exists. Moreover, the fact that the IPCC has a consensus position, despite so much cited disagreement, suggests something worse, that it is potentially fabricated. The Petition Project, for example, claims:

> The United Nations IPCC also publishes a research review in the form of a voluminous, occasionally-updated report on the subject of climate change, which the United Nations asserts is "authored" by approximately 600 scientists. These "authors" are not, however – as is ordinarily the custom in science – permitted power of approval the published review of which they are putative authors. They are permitted to comment on the draft text, but the final text neither conforms to nor includes many of their comments. The final text conforms instead to the United Nations objective of building support for world taxation and rationing of industrially-useful energy. (Petition Project[59])

The suggestion here is that the IPCC's review process is not genuine, and that the final draft of the assessment report is written in a manner that helps the United Nations get leverage to implement global taxation and rationing of energy. Given this, the IPCC should not be trusted as an expert body on climate change, and no climate policies should be implemented because it would unreasonably hurt the economy (i.e., fossil fuel industries). Now that we have a sense of how some skeptics approach criticizing the IPCC's consensus position, I next turn attention to Naomi Oreskes's defense of the IPCC.

[59] http://www.petitionproject.org/review_article.php

III Responding to the Skeptics: Oreskes's Strategy

One well-known defender of climate science research is the trained-geologist turned historian of science Naomi Oreskes. Described as a "lightning rod" in the headline of a *New York Times* profile piece in 2015, Oreskes's defense of climate science research has touched both academic and public audiences.[60] For example, her popular book *Merchants of Doubt* (2010) co-authored with Erik M. Conway reveals the connections between the tobacco and fossil fuel industries, and their aims to manufacture doubt about scientific research that threatens their businesses.[61] In this section, I focus on two papers Oreskes published in 2004: "Science and Public Policy: What's Proof Got to Do with It?" and "The Scientific Consensus on Climate Change." The latter, in particular, has received much attention from academics and the public, including skeptics. Indeed, this article is one of the reasons why there is so much discussion about the consensus of experts when it comes to anthropogenic climate change.[62] I focus on both papers to show Oreskes's broader argumentative strategy.

In "Science and Public Policy: What's Proof Got to Do with It?" (2004a), Oreskes argues that scientific knowledge cannot be assessed using the standard of proof. She writes: "the idea that science ever could provide proof upon which to base policy is a misunderstanding (or misrepresentation) of science, and therefore of the role that science ever could play in policy. In all but the most trivial cases, science does not produce logically indisputable proofs about the natural world" (2004a, 369). Scientific practice is a social and collective endeavor, and the knowledge produced is contingent on the judgments and methodologies of the community's current practices. There is no atemporal standard to ensure that evidence for scientific knowledge offers 'logically indisputable proof.' Even if experts generally agree, they can still disagree about specific details or anomalies (2004a, 389). So rather than wait for "logically indisputable proofs," Oreskes argues that the consensus of expert scientists is the relevant criterion for determining trustworthy scientific knowledge. She writes:

> [T]he appropriate standard for judging science is neither proof, nor certainty, not unanimity, but a broad and firm consensus of the relevant experts in the field. The reason is simply this: Scientific knowledge is the intellectual and social consensus of affiliated experts based on the weight of available empirical evidence and evaluated according to accepted methodologies. If

---

[60] The piece describes her as follows:

> Dr. Oreskes is fast becoming one of the biggest names on climate science - not as a climatologist, but as a defender who uses the tools of historical scholarship to counter what she sees as ideologically motivated attacks on the field. [...] Formally, she is a historian of science. Informally, this diminutive woman has become a boxer, throwing herself into a very public arena that many career-minded climate scientists try to avoid. (*NYT* 2015)

See: https://www.nytimes.com/2015/06/16/science/naomi-oreskes-a-lightning-rod-in-a-changing-climate.html

[61] Also see Oreskes (2019).

[62] According to *Science*'s AltMetrics article usage, the paper's abstract was accessed online 1878 times, the full article accessed 6687 times, and .pdf accessed 1577 times – all in January 2020 alone.
See: https://science.sciencemag.org/content/306/5702/1686/tab-article-info

> we feel that a policy question deserves to be informed by scientific knowledge, then we have no choice but to ask, what is the consensus of experts on this matter? If there is no consensus of experts – as was the case among earth scientists about moving continents before the late 1960s – then we have a case for more research. If there is a consensus of experts – as there is today over the reality of anthropogenic climate change (Oreskes 2004[b])—then we have a case for moving forward with relevant action. (2004a, 372)

The consensus of experts is our best gauge for determining whether scientific knowledge, as it currently stands, is trustworthy. However, it does not offer indisputable proof, and should be subject to continued critique and revision if further evidence becomes available. Regarding anthropogenic climate change then: is there a consensus of experts? Oreskes answers 'yes'.

In "The Scientific Consensus on Climate Change"[63] (2004b), Oreskes offers an aggregative assessment of the IPCC's consensus position. She starts by pointing out that "[t]he scientific consensus is clearly expressed in the reports of the [IPCC]" and the "IPCC states unequivocally that the consensus of scientific opinion is that Earth's climate is being affected by human activities [...]" (2004a, 1686).[64] But she notes that one might still be skeptical of the conclusion reached by the IPCC. It is possible that dissenting voices were "downplayed" or ignored during the process of deliberation.[65] Oreskes therefore tests this skeptical view by analyzing the abstracts of 928 peer-reviewed articles in scientific journals using the keywords "climate change" between 1993 and 2003. If dissenting opinions were indeed downplayed or silenced during the IPCC group's discussions and deliberations, one place it would likely emerge is in the scientific peer-reviewed literature. However, the result of the aggregative survey showed that none of the papers rejected the IPCC's consensus position. Oreskes writes: "Of all the papers, 75% fell into the first three categories, either explicitly or implicitly accepting the consensus view; 25% dealt with methods or paleoclimate, taking no position on current anthropogenic climate change. Remarkably, none of the papers disagreed with the [IPCC's] consensus position" (2004b, 1686). This result, Oreskes argues, legitimates the IPCC's consensus position. It is corroborated by an independent assessment of a decade's worth of peer-reviewed literature. Indeed, the survey shows there is a "broad and firm consensus" on the reality of anthropogenic climate change.

So how does Oreskes's argumentative strategy affect skeptics' appeal to the unanimity criterion? The skeptic – perhaps intentionally – does not specify how much dissent warrants doubt about anthropogenic climate change. If they require "logically indisputable proof," so no dissent, then it is an impossible standard. As Oreskes notes, there is no atemporal standard to ensure that evidence for scientific claims is absolute. Alternatively, if the skeptic requires a majority consensus or something

[63] See Cook et al (2013) for a more recent survey, which supports Oreskes's survey.
[64] She also cites the conclusion of the IPCC's report: "[h]uman activities [...] are modifying the concentration of atmospheric constituents ... that absorb or scatter radiant energy [...] [Most] of the observed warming over the last 50 years is likely to have been due to the increase in greenhouse gas concentrations" (2004a, 1686).
[65] This is exactly what proponents of the Petition Project mentioned earlier claim.

close to unanimity, while leaving some room for dissent, then Oreskes's aggregative survey offers evidence for just that. Recall that none of the papers surveyed undermine or reject anthropogenic climate change.[66] Oreskes's arguments thus appear to offer a way to adequately address skeptics' doubts. First, we can reject the unanimity criterion for aggregative consensus as an appropriate standard if it amounts to logically indisputable proof. Second, if the unanimity criterion for aggregative consensus amounts to majority or even near complete consensus, then we can grant it as a criterion, and then demonstrate through Oreskes's aggregative survey that it can, indeed, be met. In the next section, I look more closely at Oreskes's argument, and point out a problem.

IV A Factor Not Considered

While Oreskes's efforts to take on the skeptics is commendable, I think there are downsides to defending the IPCC's consensus position (and their assessment reports) in the way that she has. First, an important factor not acknowledged in Oreskes's approach is that many of the papers surveyed are jointly authored. In the first year of the survey (1993) the ratio between single authored and jointly authored papers was 70 to 126, respectively.[67] Jointly authored papers outnumbered single authored papers substantially. I only present data from 1993 here, but there are good reasons to assume co-authored papers likely increased over the course of the remaining nine years of the survey. One reason is the increase in co-authorship in scientific publications generally, which is now the norm (Morrison et al 2003; NSF 2018). Another reason is that climate science research is often collaborative and interdisciplinary requiring multiple contributors with different expertise. Even if a static ratio between single authored and jointly authored papers is assumed, based on the 1993 data, the former would still out-number the latter substantially. Given this, there is good reason to think co-authored papers make up a significant portion of the publications considered in Oreskes's survey. This matters because single-authored papers are produced by one person who is expected to be competent about the contents of their paper. Co-authored papers, on the other hand, are the product of a team effort. The group position expressed in an article is a result of a collaborative effort. More specifically, collaborative groups divide cognitive labor with respect to expertise, and members rely on each other both to make individual contributions and to fulfill the obligations expected of them as a collaborator (Andersen and Wagenknecht 2013; Hardwig 1985; Wray 2007). An important consequence of collaborative efforts is that no single group member may be in a position to assess or carry out all of the components of a particular project alone. Depending on the sort of group under consideration, individual scientists can choose to collaborate for a

[66] By contrast to the skeptics, Oreskes's discussion of expert consensus pertains to the "relevant experts in the field" of climate science (2004a, 372). Simply being a scientist (e.g., Fred Singer) does not warrant the right kind of membership to the climate science community.
[67] 196 articles were surveyed in total for 1993.
http://iopscience.iop.org/1748-9326/8/2/024024/media/erl460291datafile.txt

variety of reasons, which might include: the need for others' expertise that they do not have, the need for additional collaborators because of a project's large scale, or perhaps for the purposes of including a diverse set of perspectives. In addition, there are collaborations that involve researchers at various stages of their academic career trajectories (e.g., PI, postdoc, graduate student, undergraduate researcher). In these contexts, some researchers will have more experience and more expertise than other members in the group, and junior collaborators will defer to the senior collaborators to make judgment calls on issues for which they possess little or no experience. Whatever the reason, jointly authored papers are different from single-authored papers because they are often the product of collaborative efforts.

However, Oreskes does not explicitly acknowledge that many of the papers surveyed are jointly authored. This is most telling in her response to skeptics in which she effectively accepts aggregative consensus and the threshold of near or complete unanimity as the way to corroborate the IPCC's consensus position. Oreskes's survey counts the positions asserted in published articles on anthropogenic climate change to show "that scientists publishing in peer-reviewed literature agree with the IPCC […]" (Oreskes 2004a, 1686). But aggregative consensus, with its focus on individual scientists' judgments, is ill-suited to making sense of group dynamics in collaboration, such as divisions of cognitive labor and deference to others' expertise. It does not account for the collaborative nature of such groups – it notably leaves it out. If my analysis is right, then Oreskes's reliance on a notion of aggregative consensus in the context of this public debate seems problematic for a couple of reasons. First, the consensus that Oreskes establishes in her survey might not be convincing because it is unclear whether group positions expressed in the joint publications surveyed can be straightforwardly aggregated given the collaborative nature of these groups. If for many of these jointly authored publications group members rely on the work of their fellow collaborators, and thus are frequently not in a position to judge all of the paper's contents, then it seems an open question whether group positions can be aggregated and understood in the same manner as aggregating the positions of individuals. Second, a downside to granting aggregative consensus as an appropriate measure of expert consensus in this debate is that skeptics and/or those who stand to benefit from weak climate policies can always cite more dissenting "scientists." And it seems that they have no shortage of scientists (and public figures such as politicians) who are willing to fight on their behalf. Defenders of climate science, including Oreskes and climate scientists, will face an ongoing battle of demonstrating that there is not the level of dissent among relevant experts that the skeptics claim. Relatedly, granting the appropriateness of aggregative consensus in this debate also sets up the public's expectation, for better or for worse, that this is what expert consensus amounts to.[68] But I think there is room to question whether this is the right approach to respond to skeptics, which inevitably will affect the public's perception about the issue.

[68] As noted in Chapter 1 though, there isn't one way to achieve expert consensus, and much will depend on the purposes for achieving consensus.

In the next section, I outline an alternative response to skeptics, one that rejects aggregative consensus as the appropriate category of consensus to consider. I argue that even though the IPCC asserts a consensus position on the status of climate change, it is not aggregative consensus, but arguably a collaborative consensus position. If I am right, then both the skeptics and Oreskes have misplaced their attention on aggregative consensus.

V The IPCC's Consensus Position

In this section, I show the IPCC, as a group, is engaged in collaboration and that their group position on anthropogenic climate change should be understood as a collaboratively produced consensus position.

The IPCC is comprised of many climate scientists, divided into three Working Groups (I, II, and III), a task force, a task group, and other units (see Figure 2). Each Working Group covers a particular area of research with main topics and they produce a report that reflects the most current research.

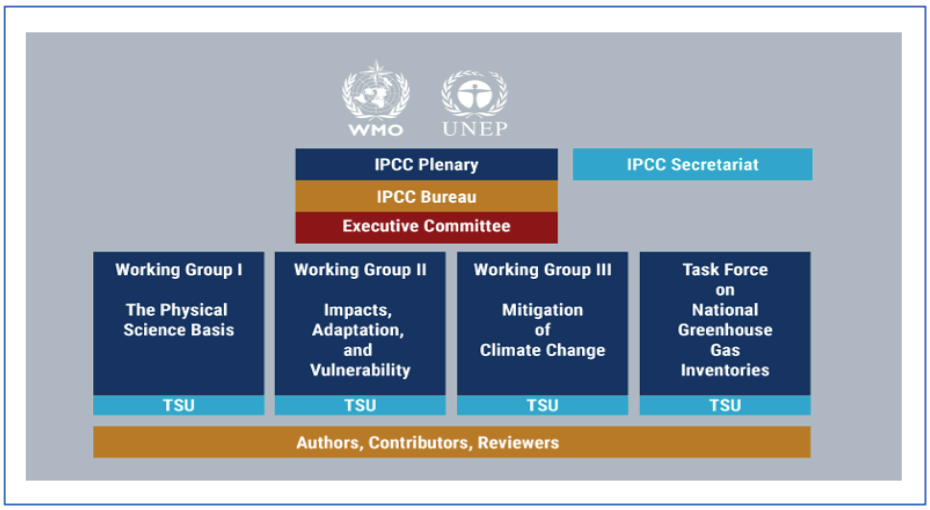

Figure 2: The IPCC's organizational structure (IPCC 2020a).

Working Group I, for example, "assesses the physical scientific aspects of the climate system and climate change" (IPCC 2020c). Some topics include changes in greenhouse gases in the atmosphere, climate models, and climate projections (IPCC 2020c). The group is comprised of many climate scientists with differing expertise, and no single member is in a position to assess the contents of all three Working Groups. The group divides cognitive labor among the three Working Groups, and members of one Working Group will defer to others on matters pertaining to another Working Group. As well, there is a division of cognitive labor within each Working Group. For example, "Lead Authors" of a given report can seek out the expertise of "Contributing Lead Authors", which they do not have, to prepare technical

information for the report (e.g., graph or data).[69] Given this organizational structure, the IPCC, as a group, operates like a collaboration.

In addition, a central feature of IPCC's assessment reports is the drafting and review process. The IPCC describes their process as follows:

> IPCC reports undergo multiple rounds of drafting and review to ensure that they are comprehensive and objective and produced in an open and transparent way. Thousands of other experts contribute to the reports by acting as reviewers, ensuring the reports reflect the full range of views in the scientific community. Teams of Review Editors provide a thorough monitoring mechanism for making sure that review comments are addressed. (IPCC 2013b)

The reports are subject to multiple rounds of revision to include comments and feedback from thousands of expert reviewers to ensure diverse perspectives are considered. Authors and reviewers participate in multiple rounds of exchange that involve deliberation as well as negotiation since authors must address review comments and reviewers comment on multiple drafts of the report. In this sense, the IPCC report importantly does not reflect an aggregation of authors' views about the main topics that they are tasked to write on. Rather, the review process is meant to help the group establish how best to present the latest findings on climate change. So while the scientists authoring the report stand by its contents, they may not individually agree or even know everything in it. The report strikes a negotiated compromise among the participants of any given chapter and overall Working Group report. Gilbert's joint commitment model is especially helpful here. The model can help explain that members can agree to let the product stand as the group view, even if they have different reasons for doing so. Moreover, the group position is more comprehensive than what each individual member can contribute because of the multiple rounds of revisions requiring that comments from expert reviewers be addressed. The production and publication of these assessment reports every 5-7 years could not take place without there being something like a joint commitment (in Gilbert's sense) by participants to collaborate and produce it.[70] This commitment binds participants to undergo the group's procedures and practices to complete the reports.

I have discussed two notable features about the IPCC group. One is that they operate like a collaboration, given their organizational structure and sophisticated division of cognitive labor both within and across the different Working Groups. The other is that the review process of their assessment reports undergo multiple rounds of revision with feedback from many reviewers. Taken together, these features provide reasons to understand the IPCC's consensus position on anthropogenic climate change (drawn from the assessment report) as a collaboratively produced one. Gilbert's joint commitment model in particular has resources to explain how there can be a group consensus position even though there are various working groups, a variety of different experts, and areas of disagreement: members have agreed to let the position stand as the group's. So the IPCC's consensus position is not aggregative consensus,

[69] See Chapter 5.
[70] More on Gilbert below.

but collaborative consensus achieved via something like joint commitment. The evidence in support of climate change is, of course, a central component of the report. But the IPCC's joint products also centrally depend on the willingness of group members to engage in a joint endeavor where they reliably depend on each other for the group's success. More specifically, the IPCC's consensus position and their assessment report are dependent on the different contributions of its members, and no individual alone could have done this. In this sense, aggregative consensus is not a relevant marker for assessing the IPCC's consensus position for it presumes that the IPCC achieved consensus by simply aggregating the beliefs of participating climate scientists. But as I have shown this is not how the group operates.

If I am right that we should understand the IPCC group as engaging in collaboration and that their joint product, particularly their public consensus position on anthropogenic climate change (and their assessment reports), amount to collaborative consensus, then we have a different option to respond to skeptics. That is, to reject aggregative consensus as the appropriate way to evaluate the IPCC's consensus position because they have produced collaborative consensus. Every time a skeptic demands that among climate scientists there must be near or complete unanimous aggregative consensus, defenders of climate science ought not fall into the trap of trying to provide evidence for this. As I mentioned earlier, if we grant this demand, we will continually be stuck in a cycle where we have to keep responding to "experts" who disagree. Rather, we should reject the demand, and point out that no single climate scientist (who publishes in peer-reviewed climate science journals) can know everything there is to know about the issue. So while I commend Oreskes for her efforts, I think her comprehensive survey of peer-review literature to corroborate the IPCC's consensus position, which so many climate change defenders cite, actually does more harm than good because it wrongly places importance on aggregative consensus.

Conclusion

In this chapter, I considered the public controversy about the IPCC's consensus position on the reality of anthropogenic climate change, focusing on the arguments of some climate change skeptics and a prominent defender of climate science, Naomi Oreskes. I pointed out that both the skeptics and Oreskes focus on aggregative consensus as the relevant type of expert scientific consensus in this debate. The skeptics challenge the IPCC's consensus position by emphasizing that there actually is no aggregative consensus. Instead, there is substantial disagreement among scientific experts. Oreskes, on the other hand, defends and corroborates the IPCC's consensus position by conducting a comprehensive survey aggregating the positions of peer-reviewed articles on climate change. Oreskes argues that not only do the authors of these articles agree with the IPCC, but also none of them challenged the status of anthropogenic climate change. The approach, however, is problematic because many of the papers surveyed are jointly authored, and it is unclear whether there is a straightforward way to aggregate collaboratively produced group positions. Moreover, I briefly analyzed the IPCC's organizational structure and their processes that help produce the assessment report. I concluded that the group operates like a

collaboration and that their consensus position is a collaboratively produced one. Given this, we should resist temptation to respond to skeptics' demand for aggregative consensus because it has limited relevance for assessing the IPCC's consensus position. And consequently, I think Oreskes's argumentative strategy is both mistaken and problematic. It sets up the public's attention to focus on aggregative consensus, which may give the perception that it is the only way to achieve expert consensus.

From here, the way forward is to consider what aspects of collaboration make collaboratively produced consensus positions trustworthy. If we can identify features that help locate well-functioning collaborations, then this may yield a way to positively defend the procedures and practices of the IPCC that helped produce their consensus position. In the next chapter, I discuss what I argue to be crucial features of scientific collaboration.

Bibliography

Andersen, Hanne. 2010. “Joint Acceptance and Scientific Change: A Case Study.” *Episteme* 7 (3): 248–65.

Andersen, Hanne, and Susan Wagenknecht. 2013. “Epistemic Dependence in Interdisciplinary Groups.” *Synthese* 190: 1881–98.

Beatty, John. 2006. “Masking Disagreement Among Experts.” *Episteme* 3 (1–2): 16.

———. 2017. “Consensus: Sometimes It Doesn’t Add Up.” In *Landscapes of Collectivity in the Life Sciences*, edited by Snait Gissus, Ehud Lamm, and Ayelet Shavit, 179–89. Cambridge, MA: MIT Press.

Beatty, John, and Alfred Moore. 2010. “Should We Aim for Consensus?” *Episteme* 7 (3): 198–214.

Boyer-Kassem, Thomas, Conor Mayo-Wilson, and Michael Weisberg, eds. 2018. *Scientific Collaboration and Collective Knowledge*. New York: Oxford University Press.

Bratman, Michael. 1999. *The Faces of Intention*. Cambridge, MA: Cambridge University Press.

Ceccarelli, Leah. 2011. “Manufactured Scientific Controversy: Science, Rhetoric, and Public Debate.” *Rhetoric & Public Affairs* 14 (2): 195–228.

Cook, John, Dana Nuccitelli, and Sarah A. Green. 2013. “Quantifying the Consensus on Anthropogenic Global Warming in the Scientific Literature.” *Environmental Research Letters* 8 (2): 024024.

Fagan, Melinda Bonnie. 2012. “Collective Scientific Knowledge: Collective Scientific Knowledge.” *Philosophy Compass* 7 (12): 821–31. https://doi.org/10.1111/j.1747-9991.2012.00528.x.

Gilbert, Margaret. 1996. *Living Together*. Landham, MD: Rowman & Littlefield Publishers.

Gillis, Justin. 2015. “Naomi Oreskes, a Lightning Rod in a Changing Climate.” *New York Times*, June 15, 2015, sec. Profiles in Science. https://www.nytimes.com/2015/06/16/science/naomi-oreskes-a-lightning-rod-in-a-changing-climate.html.

Goodwin, Jean. 2009. “The Authority of the IPCC First Assessment Report and the Manufacture of Consensus.” *Working Draft*.

Hardwig, John. 1985. “Epistemic Dependence.” *The Journal of Philosophy* 82 (7): 335–49.

Houghton, J.T., G.J. Jenkins, and J.J. Ephraums. 1990. *Climate Change: The IPCC Scientific Assessment: Report Prepared for IPCC Working Group 1*. Cambridge: Cambridge University Press.

Intemann, Kristen. 2017. “Who Needs Consensus Anyway? Addressing Manufactured Doubt and Increasing Public Trust in Climate Science.” *Public Affairs Quarterly* 31 (3): 189–208.

IPCC. 2013a. “Summary for Policy Makers.” Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom, and New York, NY USA.

———. 2013b. “IPCC Factsheet: What Is the IPCC?” https://www.ipcc.ch/site/assets/uploads/2018/02/FS_what_ipcc.pdf.

———. 2013c. “Principles Governing IPCC Work.” Batumi. https://archive.ipcc.ch/pdf/ipcc-principles/ipcc-principles.pdf.

———. 2015. “IPCC Factsheet: How Does the IPCC Review Process Work?” January 15, 2015. https://www.ipcc.ch/site/assets/uploads/2018/02/FS_review_process.pdf.

———. 2020a. “Structure of the IPCC.” IPCC. 2020. https://www.ipcc.ch/about/structure/.

———. 2020b. “The Intergovernmental Panel on Climate Change.” IPCC. 2020. https://www.ipcc.ch.

———. 2020c. “Working Groups.” IPCC. 2020. https://www.ipcc.ch/working-groups/.

Loris, Nicolas D. 2010. “How the ‘Scientific Consensus’ on Global Warming Affects American Business—and Consumers.” *The Heritage Foundation Backgrounder*, 2010, No. 2479 edition.

Miller, Boaz. 2016. “Scientific Consensus and Expert Testimony in Courts: Lessons from the Bendectin Litigation.” *Foundations of Science*. https://doi.org/10.1007/s10699-014-9373-z.

Morrison, Philip, Gill Dobbie, and Fiona J. McDonald. 2003. “Research Collaboration Among University Scientists.” *Higher Education Research & Development* 22 (3): 275–96.

National Science Foundation. 2018. “Science and Engineering Indicators 2018.” https://www.nsf.gov/statistics/2018/nsb20181/report/sections/academic-research-and-development/highlights.

Nisbet, M. C., and T Myers. n.d. “The Polls - Trends - Twenty Years of Public Opinion about Global Warming.” *Public Opinion Quarterly* 71: 444–70.

Oreskes, Naomi. 2004a. “Science and Public Policy: What’s Proof Got to Do with It?” *Environmental Science and Policy* 7: 369–83.
———. 2004b. “The Scientific Consensus on Climate Change.” *Science* 306: 1686.
———. 2019. *Why Trust Science?* Oxford: Princeton University Press.
Oreskes, Naomi, and Erik M. Conway. 2010. *Merchants of Doubt*. New York: Bloomsbury.
Petition Project. n.d. “Global Warming Petition Project.” http://www.petitionproject.org/index.php.
Rehg, William, and Kent Staley. 2017. “‘Agreement’ in the IPCC Confidence Measure.” *Studies in History and Philosophy of Modern Physics* 57: 126–34.
Schmitt, Frederick F. 1994. “The Justification of Group Beliefs.” In *Socializing Epistemology: The Social Dimensions of Knowledge*, edited by Frederick F. Schmitt. Maryland, USA: Rowman & Littlefield Publishers.
Science. 2020. “The Scientific Consensus on Climate Change.” Article Information. Science Magazine. July 2020. https://science.sciencemag.org/content/306/5702/1686/tab-article-info.
Solomon, Miriam. 2006. “Groupthink versus the Wisdom of Crowds: The Social Epistemology of Deliberation and Dissent.” *The Southern Journal of Philosophy*, no. 44 Supplement: 28–42.
Staley, Kent. 2007. “Evidential Collaborations: Epistemic and Pragmatic Considerations in ‘Group Belief.’” *Social Epistemology* 21 (3): 15.
Tuomela, Raimo. 1992. “Group Beliefs.” *Synthese* 91 (3): 285–318.
Wray, K. Brad. 2007. “Who Has Scientific Knowledge.” *Social Epistemology* 21 (3): 337–47.

4 Scientific Collaboration: Epistemic Dependence and Social Roles

In the last chapter, I briefly looked at the organizational structure and processes that guide the production of the IPCC's work, and argued that they operate as a collaboration. Their consensus position on anthropogenic climate change should be understood as a collaboratively produced one. Skeptics are unreasonable to demand that the IPCC's consensus position ought satisfy near or complete unanimity. And Oreskes's attempt to meet the skeptics' demand in defense of the IPCC is a mistake. We need a different strategy to defend the IPCC's consensus position. In the remaining chapters of this dissertation, my aim is to identify important features of scientific collaboration, and then propose a set of working normative conditions for well-functioning scientific collaboration with these features in mind. Then I will assess the IPCC's organizational structure and processes according to these normative conditions. Overall, my position is that the degree to which a group's collaborative consensus position satisfies the normative conditions outlined indicates reasons for whether to trust them. In this chapter, I argue that an important feature of scientific collaboration has to do with collaborators' relations of dependence. In particular, I look at relations of epistemic dependence, as well as the norms and obligations associated with each collaborator's social role in the group. I make use of these relations as a crucial starting point to conceptualize normative conditions for well-functioning collaboration.

This chapter is organized into four sections. In sections I and II, I highlight two ways in which collaborators are relationally dependent on each other. First, drawing from literature concerning the social epistemology of science, I discuss relations of epistemic dependence among collaborators. Second, I discuss social roles, and how collaborators rely on them to form norms about what they can reasonably expect from fellow collaborators. I look at social roles in research collaboration and co-authorship, which show how social roles in important contemporary collaborative practices structure groups hierarchically. In section III, I argue relations of epistemic dependence and social roles interact, for example, when collaborators are working together on a co-authored paper, which impacts how members' contributions can be given more or less weight. Finally, with two significant types of dependence relations for collaboration on the table, I consider whether it makes sense to use Longino's norms for scientific knowledge inquiry to assess collaborations. I conclude, however, that Longino's norms are not a good fit because they do not include consideration about authority structures.

I Epistemic Dependence

In the last chapter, I argued that the IPCC is engaged in collaboration, and I highlighted some features to demonstrate this. These features are also common to significant types of collaborative practice in science such as research collaboration and co-authorship. One is that group members share and divide cognitive labor to realize joint goals (e.g., publishing a joint product or achieving a research goal). Relatedly, another is that many collaborators in a given group may not be in a position to competently assess all of their collaboration's scientific contents for a variety of reasons, such as the project's large scale, finite expertise, or time constraints. Finally, group consensus positions can be subject to negotiation and compromise, often geared towards realizing the group's joint goals. Taken together these features suggest that members of collaborations are importantly dependent on each other (Wray 2007). For example, they depend on each other epistemically when they do not possess the necessary experience or expertise to carry out certain tasks or when they rely on each other to provide substantive feedback. In addition, collaborators are organizationally dependent on each other since there is often a structure for sharing and dividing labor, as well as assigning credit. Note that these dependence relations can occur both in contexts where members share similar expertise (Staley 2007), and in interdisciplinary contexts where collaborators have different expertise (Andersen and Wagenknecht 2013). Understanding how some of these dependence relations work in collaboration provide a fuller picture of consensus formation by collaborative groups. My aim is to locate and make explicit some central dependence relations, which will help inform later discussions about normative conditions for scientific collaboration. I will focus on two dependence relations that I take to be central to collaborative practice in science: (1) relations of epistemic dependence, which have received attention from philosophers, and (2) relations of dependence based on the social roles collaborators occupy, which have been given less attention.

Scientists are often epistemically dependent on the work of others.[71] John Hardwig's (1985) account of epistemic dependence offers a helpful starting point. He conceptualizes epistemic dependence among scientists as similar to the relation between experts and lay people. We often think of lay people as those who rely on experts' judgments about matters they do not know. But, Hardwig argues, a similar relation can occur between scientific experts. When scientists conduct research it often, broadly speaking, relies on the work of others. That is, they build on existing results, for example, by taking them as starting points. In other contexts, they take existing results as a point of contention or disagreement from which to develop their work. One thus might think that no individual scientist is capable of conducting research without being epistemically dependent on the work of other scientists, given cognitive limitations as well as time constraints. Hardwig writes:

[71] Some philosophers have developed mathematical models of divisions of cognitive labor in science (Kitcher 1992; Weisberg and Muldoon 2009; Zollman 2010), but they do not adequately model collaborations where members work together (Muldoon 2018).

> Scientists, researchers, and scholars are, sometimes at least, knowers, and all of these knowers stand on each other's shoulders in the way expressed by the formula: *B* knows that *A* knows that *p*. These knowers could not do their work without presupposing the validity of many other inquiries which they cannot (for reasons of competence as well as time) validate for themselves. […] It would, moreover, be impossible for anyone to get to the research front in, say, physics or psychology, if he relied only on the results of his own inquiry or insisted on assessing for himself the evidence behind all the beliefs he accepts in his field. Thus, if scientists, researchers, and scholars are knowers, the layman-expert relationship is also present *within* the structure of knowledge, and the expert is an expert partially because he so often takes the role of the layman *within his own field*. (1985, 345-346; original emphasis)

On Hardwig's analysis, scientific experts can themselves operate like laypeople in relation to other experts in a few different ways. One is that they do not have the capacity or relevant expertise to conduct certain work. Another is that they rely on existing work from which they build their own work. One may, however, be reluctant to grant this layman-expert relationship when a researcher relies on established work, if they have the capacity to understand and, at least in theory, maybe even conduct it. For my purposes though, the important point to acknowledge is that experts are epistemically dependent on the work of others when they pursue their own research. One need not think that this laymen-expert relationship applies in all cases. In interdisciplinary collaborations where specializations are vastly different one might grant this laymen-expert relationship, but perhaps in collaborations where members share similar training we may not grant this relationship. Either way, of importance is that scientists do indeed rely on the work of others. Relations of epistemic dependence can, in short, vary in degrees.

For my purposes, I emphasize that relations of epistemic dependence are essential in teamwork contexts, such as collaborations resulting in joint publication. Hardwig discusses a paper from The American Physical Society on particle physics published in 1983, comprised of 99 co-authors (1985, 346). Here no single author is alone capable of conducting the research on charm events and charm particles reported in the article for at least two reasons. Even though these co-authors are particle physicists, they have different specializations, and so they rely on each other with respect to specialized topics with which they are less familiar. And without such a joint effort, it would take an individual far too long to conduct the research, which spans multiple research sites. Relations of epistemic dependence are also relevant when considering interdisciplinary research collaborations. In such cases, members often defer to each other on matters for which they are not the requisite authorities. As Hardwig notes, specialization and teamwork are "inescapable features of much modern knowledge acquisition" (1991, 696).

Hanne Andersen and Susann Wagenknecht (2013) build on Hardwig's account of epistemic dependence and distinguish between "different patterns of interaction," which they call "unilateral epistemic dependence" and "multilateral epistemic dependence." Unilateral epistemic dependence occurs when individual scientists trust and rely on the informational testimony of other scientists. They

combine this testimony with knowledge they possess to make new inferences (2013, 1884). For example, a scientist extracts informational testimony from a peer-reviewed publication or a conference presentation to make a new inference that combines their own research. Given the unilateral, "chain-like" pattern of interaction, the scientists involved need not be collaborators. "Multilateral epistemic dependence" takes place when scientists are "*mutually* dependent on each other and therefore develop multilateral relations of trust to each other" (2013, 1885). Hardwig's example of the particle physics paper with 99 co-authors would count as one example of multilateral epistemic dependence. Unilateral epistemic dependence captures one-way relations of epistemic dependence, which may or may not involve collaboration; while multilateral epistemic dependence captures epistemic dependence that does involve collaboration.

Andersen and Wagenknecht also consider how do collaborators make new inferences, using testimony provided by others in the group, when they do not share the same expertise. It seems that one would need some understanding of claims to be able to integrate them and then to infer a new one. Andersen and Wagenknecht write: "[…] in actual scientific research, usually quite a lot of knowledge about the scientific claim that *n* and about the scientific claim that *o* will be necessary in order to combine them as premises leading to the conclusion that *p*" (2013, 1886). They call this the problem of "cognitive integration." The authors offer a potential solution by drawing from science studies and social science research (e.g., Collins and Evans 2003; Collins et al 2007). They suggest that mutual learning, as well as some shared mental models (or at least some aspects of one), among collaborators in interdisciplinary groups is needed.

Andersen and Wagenknecht draw on empirical research by Frederick Rossini and Alan Porter (1979), which offers four ideal models that show how cognitive integration in interdisciplinary groups can occur. One model is "integration by leader," which involves an individual scientist taking the lead role of interacting with and then integrating contributions from each collaborator (1979, 1888).[72] The leader takes on the task of cognitive integration. A second model is "common group learning," which involves collaborators learning the same content together until it becomes common knowledge. As a consequence, all members are able to make the same inferences about the learned content, so cognitive integration is a joint endeavor (1979, 1889-1890).[73] A third model is "negotiation among experts." Cognitive labor is distributed based on members' expertise, and the task of cognitive integration occurs iteratively through negotiation among members (1979, 1890). Finally, there is "modelling," which occurs when information from different experts is "fed" into a chosen model that then integrates and organizes

[72] Note that this model maps onto their notion of "unilateral epistemic dependence" mentioned above. But given that they are considering interdisciplinary groups, this model does involve collaboration between group members and the leader integrating the knowledge. Andersen and Wagenknecht also note that this approach is limited in the research it can cover, given the amount of work that is required of the leader,

[73] This approach is also limited because it requires everyone in the group to invest cognitive resources to learn content from a variety of disciplines, which may be too demanding for individual scientists.

the information (1979, 1891). The chosen model completes the task of cognitive integration. These ideal models offer a window into some potential solutions that resolve the problem of cognitive integration in interdisciplinary collaboration, where members' expertise do not overlap.

So while Hardwig's account of epistemic dependence is a good starting point for considering relations of epistemic dependence, Andersen and Wagenknecht show that there can be different types of epistemic dependence – unilateral and multilateral epistemic dependence. They also point out that any account of epistemic dependence would need to explain how cognitive integration occurs not only for collaborations where members share similar or overlapping expertise but also for interdisciplinary collaborations, which are more complicated. Additional socio-cognitive models like Rossini and Porter's capture different patterns of interaction and different types of cognitive integration that occur among collaborative groups. I have made use of some of the literature on epistemic dependence to help clarify how we should understand this relation and how it operates in collaboration. But as I mentioned at the beginning, collaborators are dependent on each other in other ways too. Next, I introduce social roles as another important dependence relation.

II Social Roles

Another type of dependence relation in scientific collaboration that has not received sufficient attention is members' assigned social roles in the group. The PI position, for instance, is often reserved for someone who has completed their doctorate, possesses a good amount of research experience, and has secured a research grant. Social roles help outline (both tacitly and explicitly) some of the central tasks and duties expected of each collaborator, which is common knowledge to group members. For example, a PI's obligations in a collaboration are quite different from a graduate student's obligations. In addition, collaborators depend on assigned social roles to help form norms about what each person can reasonably contribute to the group. Finally, a distinctive feature of social roles that I would like to emphasize is that they can confer differential degrees of authority, which affects the group's organizational structure. Note that this is not an anomaly in scientific collaboration – it is quite common, as I will discuss. My discussion will focus primarily on hierarchically organized collaborations where some group members have greater authority and others have less, in relation to one another. I will look at two examples from collaborative practice. The first example will highlight the various social roles within a more standard case of collaboration – an engineering research center. Then I turn attention to social roles in the context of co-authorship. These examples are inevitably related since co-authorship is often an explicit goal of collaborative groups.[74] Publications are, for better or for worse, one of the primary

[74] Thanks to Andrea Woody for encouraging me to clarify the relationship between research collaborations and co-authorship. Though, I still have not said enough and hope to further articulate this relation in future work.

ways in which the success or failure of a collaboration is measured. They are used both to disseminate work to gain credit in research communities, and to show funding agencies that progress has been made.

Engineering Research Centers

One frequently discussed example is the CERN-based Large Hadron Collider collaboration, which is comprised of two different sites, each with fifteen hundred to two thousand physicists (Galison 2003; Huebner et al 2018). The majority of collaborators at places like CERN share the same disciplinary background, though many specialize in niche areas of research. Here though, I want to highlight another type of research collaboration that has received a lot of funding recently: NSF engineering research centers that pursue cutting-edge medical research. One exemplar is the NSF's Center for Neurotechnology (CNT), headquartered at the University of Washington in partnership with the Massachusetts Institute of Technology and San Diego State University, as well as summer exchange program with the University of Freiburg in Germany.[75] The CNT aims to produce engineered devices that read out from the brain to control an external device such as a computer cursor or speller, a wheelchair, or a prosthetic device.

The CNT is comprised of members who are assigned a variety of social roles: co-director, thrust leader (or PI), researcher, postdoc researcher, graduate student researcher, neuroethics fellow, undergraduate researcher, staff, and industry members. These different roles come with certain obligations and expectations. The co-directors are in charge of overseeing the CNT's research aims, overall activities, and progress. They are also responsible for communication with NSF site visitors to renew the Center's funding each year (for up to ten years). Thrust leaders are each in charge of overseeing and developing one of the Center's research cores, which include: experimental neuroscience, computational neuroscience, communication and interface, and neuroethics. Responsibilities include conducting research and ensuring that their group is able to report on "deliverables" such as experiments, conference presentations, and both individual and co-authored publications. Researchers and postdocs, in consultation with PIs, take on projects related to their areas of expertise and often in collaboration with other lab members. Graduate and undergraduate students are placed on projects that would help develop their research interests, while gaining practical research experience and guidance from more experienced lab members. NSF engineering research centers like the CNT could not function without members being assigned different roles, each of which have specific expectations and obligations.

Also note the hierarchical structure of the CNT: different social roles have differential degrees of authority. The co-directors arguably have the most authority since they can make executive decisions for the group, such as those related to funding allocation. They also oversee how the CNT presents itself to NSF site visitors and make recommendations to members' site visit presentations. Thrust Leaders

[75] See: http://www.csne-erc.org

oversee their own research lab but in consultation with the co-directors' guidance (especially when it is time to report to the NSF). Thrust Leaders also supervise and guide the work of postdocs, graduate students, and undergraduate students. Postdocs have more credentials than graduate students and undergraduate researchers, and they are expected to take on a more authoritative role, if they collaborate with junior members on a publication. While graduate students and undergraduate researchers have the least authority in the group, given their experience. Nevertheless, their participation is key for skill-building and development.

Co-Authorship

I also want to draw attention to another collaborative practice that structures groups according to social roles: co-authorship. Co-authorship is an important component of collaborations like the CNT discussed above since publishing research is a primary way to indicate progress to funding agencies. Co-authorship collaborations can form in variety of ways. Some are formed only when a collaboration has interesting results to report, while others are formed precisely with the goal of generating a publication. Sometimes subgroups within labs (like the research thrusts mentioned above) might pursue a related but slightly different topic from the focus of their lab's to produce their own co-authored paper. Whatever the initial intent, publications are tied to credit for intellectual work and future promotions like academic tenure. In general, co-authorship is a central collaborative practice in science, and it is now the norm and not the exception.[76]

Co-authors are typically organized and the listed according social roles that include lead author, primary investigator (PI), as well as other contributing authors. The "lead author," listed at the very front, often assumes the obligation of drafting the paper. They are also often the designated person that submits the paper to the target journal and corresponds with the journal's editor(s). If there is more than one lead author, this is specified with an asterisk or a typographical symbol, and an explanatory note (Venkatraman 2010). Depending on the lead author's career stage and their research discipline, the research project from which the paper emerged will likely come from a funded project that the PI secured, as seen in biology and biomedicine.[77] Lead author roles are highly coveted on a researcher's CV because they are attributed a high degree of credit for the paper, which an external body may use for hiring or tenure decisions. This is especially important for junior scholars who need to build a list of publications. Furthermore, it is the lead author's name that is referenced in citations – e.g., lead author et al. (XXXX). The individual in the "PI" role normally oversees the research project from which the paper emerged. If they also happen to be a lead author, then they are also listed at the front. Otherwise, they are usually listed last. Finally, all other "co-authors" are listed either after the lead author or between

[76] See Introduction Chapter.
[77] Thanks to Carole Lee for this point.

the lead author and the PI, in order of contributions to the paper. The Contributor Roles Taxonomy, or CRediT model[78], for example, offers a helpful and detailed taxonomy of contributor roles according to contribution, which accounts for the many different but important roles collaborators might occupy beyond the lead author and PI roles (CASRAI CRediT 2014). Such roles include data curation, formal analysis, and investigation among others.[79] Of importance here is that these social roles outline a particular set of expectations and obligations for each co-author that group members can in turn reasonably depend on each other to fulfill.

Social roles in co-authorship practices also confer differential degrees of authority. In the PI role one often has the most authority. Their suggestions can be weighted more heavily than, say, a graduate student or postdoc co-author (who is not a lead author). They also often have the most research and professional experience, and likely hold a secure faculty position. In addition, the project's contents and sources of funding, as mentioned, likely comes from grants that the PI secured. PIs exert authority over junior members in part because the paper's contents will be attributed to their research grant. On the other hand, lead authors who conduct significant portions of research, as well as writing, are often grad students or postdocs working in the labs of the PI. They work under the direction of the PI, and they share their ideas in consultation with the PI and the other co-authors.

The different social roles described in co-authorship practice, however, are not always strictly enforced, and much will depend on disciplinary norms. Research in biomedicine leans more toward individualizing author contributions while large scale collaborations in physics treat authorship more like "corporate bodies" (Biagioli 2003; CASRAI CRediT 2014; ICJME 2020; Wray 2018). The journal *Science*, which accepts submission from any science field, endorses the CRediT model mentioned above, which distinguishes between different contributor roles.[80] On the other hand, some journals do not require any particular policy on how co-authors should be listed. *Nature*, for example, states:

> Nature Research journals do not require all authors of a research paper to sign the letter of submission, nor do they impose an order on the list of authors. Submission to a Nature Research journal is taken by the journal to mean that all the listed authors have agreed to all of the contents, including the author list and the author contribution statements. (*Nature* 2019)

One reason for not imposing an order on the list of authors is to make room for large scale collaborations like the CERN research center (Galison 2003). Some of these collaborations actively discourage listing authors in any particular order, except alphabetically, to preserve the inclusion of group members and to preserve group unity.[81] As Peter Galison describes it: "The collaboration must function with sufficient unity for its name to stand for something" (2003, 336). Even so, the paper itself is not written by all

[78] Thanks to Carole Lee for alerting me this model, as well as the ICJME (2020) authorship guidelines.
[79] In Chapter 5, I will discuss this model in more detail.
[80] See: https://www.sciencemag.org/authors/science-journals-editorial-policies
[81] In some cases now, *only* the collaboration is listed on the publication, while the full list of alphabetized authors is listed elsewhere (see Wray 2018, 125).

collaborators. Someone or some subset of people from the collaboration write the paper, and they are designated to the "presenter" role (2003, 334). Notice that the presenter role, too, comes with certain norms and expectations – for example, presenters cannot claim that the paper's results are theirs and *not* the collaboration's. Presenters are in charge of synthesizing and explaining results that were collectively produced by the collaboration, which no individual could complete independently.

The practice of alphabetizing author names on publications by large-scale collaborations, however, is not always endorsed by members in these collaborations. Galison, for example, cites OPAL at CERN's lab which is comprised of some 2000 physicists. OPAL conducted a survey asking members about their co-authorship practices. Galison reports that "[…] some seventy-five percent of their respondents disliked publication habits, largely because it damaged the possibility of career advancement and of receiving credit for their work, and some sixty-seven percent wanted change" (2003, 341). Alphabetized publications make it harder for individual researchers to receive credit for their work, and evaluations, as a result, become more heavily based on the recommendations of others (2003, 341).[82] So while it may seem like large-scale collaborations make the practice of listing co-authors in a particular order trivial, Galison's analysis shows that they are in tension with the aim of assigning credit to researchers. Moreover, the ATLAS team, which worked on one of CERN's Large Hadron Collaboration detectors, introduced a new type of publication called the "scientific note," which allows an individual or small team of individuals to have their work undergo standard refereeing protocols (2003, 342). Scientific notes do not replace collaborations listing authors on some publications alphabetically, but they do make room for individuals and small teams to highlight their specific contributions.

Taking Stock

I have discussed two important contexts where social roles are central: research collaborations and co-authorship. Note that social roles in both contexts can overlap. In particular, social roles in co-authorship practices can reflect how a research collaboration is organized. The PI listed at the end of a co-authored publication is usually also the PI in a research collaboration who secured the funding grant. And the Lead author role usually signals that that person is a more junior scholar, such as a graduate student or a postdoc. In addition, the placement of co-authors' names on jointly authored papers signal what sort of credit to attribute them. Large scale collaborations are perhaps an exception but even there we see some recent changes to accommodate co-authorship norms to properly credit individual researchers.[83]

[82] On external perceptions, Galison writes: "In 1988, the Department of Energy committee uncomfortably contemplated the consequences of being inclusive in the author lists, as such endless rosters gave physicists from other fields the 'impression that all individuality is submerged in high energy physics', a not irrelevant image when questions of hiring, tenure, and even field support arose" (2003, 341).
[83] As I mentioned, I think more research is needed on this relationship between research collaboration and co-authorship, and the potential benefits and tensions between them.

I suggested at the beginning of the section that social roles are integral to collaborative practices as a dependence relation. In what sense then do collaborators depend on each other with respect each's social roles? Social roles designate obligations and responsibilities for individual collaborators, which other members reasonably take to be the norm. A collaboration cannot function efficiently without members doing their part, and each individual member depends and expects fellow group members to do their share. For instance, the director of a research center is responsible for reporting the center's progress to funding agencies. No one could reasonably expect the lab technician or a graduate student to be responsible for the director's obligations. Thus, collaborators depend on each other, not just to contribute their expertise, but also to realize the various duties and obligations associated with their assigned social role for the group to function well. In addition, I highlighted how social roles can confer differential degrees of authority. For example, a PI's suggestions on a paper will confer greater authority than an undergraduate researcher's suggestions. Social roles alter the dynamics of interactions among members by giving more authority to some and less to others. This also means the set of expectations and obligations for each member can vary, depending on the social role that they occupy. Those with more authority will have more responsibilities than those with less authority. Measures for culpability will vary along this scale too. From the external perspective of authoritative bodies like journal editorial boards and funding agencies, social roles are also a means of assigning credit and blame: they aim to ensure that contributions are given due credit and that mistakes are accurately attributed. The lead author of a paper or the PI will be more responsible for imprecise reporting than the undergraduate co-author who helped run the experiment.

My aim in this section was to show that in practice collaborators can and do occupy various social roles with differential authority focusing on research collaboration and co-authorship. Through these examples I also emphasized how collaborators depend on one another according to the norms and expectations associated with each's assigned social role. I have now discussed two important ways in which collaborators are relationally dependent one another – epistemically and through social roles. In the next section, I discuss an implication when these dependence relations are considered together in collaborations.

III Interaction between Epistemic Dependence and Social Roles

Each of the dependence relations structure collaborative groups by prioritizing different features – expertise for epistemic dependence and authority for social roles, respectively. In this section, I will discuss how these dependence relations also interact and impact each other. Consider a collaboration that produces a co-authored paper involving a unilateral epistemic dependence relation, where a PI and other co-authors rely on a postdoc's expertise. Assume the postdoc's dissertation is on the same topic as the group's paper, which makes the postdoc the designated expert. For this reason, they are also assigned the role of lead writer by the PI. However, the PI considers it important to be able to give substantive

feedback on the paper, even though the postdoc technically has the most relevant expertise. The PI believes their comments can help improve the chances of the paper being accepted for publication. The PI can insert their authority this way because their role confers them a greater degree of authority over other group members, including the postdoc, in virtue of their research credentials and broad academic professional experience. This does not mean they could dictate that the postdoc and other group members fabricate research or write false content.[84] But they can make calls that override the postdoc's expertise, and for which the postdoc will likely concede after discussion, if the PI insists – e.g., the strength of the paper's general claim or which research limitations to highlight. If expertise were considered central, then the group would be organized according to expertise, with the postdoc's knowledge valued most. However, in this case, the PI has decided that while the postdoc can generally be in charge of the epistemic content, they may occasionally intervene and make recommendations. This is a case in which epistemic dependence and social roles share an interactive relation: expertise is important, but someone with authority can provide comments that someone with technically more relevant expertise must take seriously. Authoritative social roles can alter or impact the effects of epistemic dependence relations.

A similar point can be made for collaborations that involve multilateral epistemic dependence relations, where members – especially in interdisciplinary groups – are mutually dependent on each other's expertise. If one focuses on epistemic dependence alone, then one may assume that cognitive integration in the group occurs democratically, since everyone is mutually dependent on each other's knowledge. One might assume that Rossini and Porter's four ideal models for cognitive integration in interdisciplinary groups (integration by leader, common group learning, negotiation, and modelling), would be decided by group members (e.g., through group deliberation and then perhaps a vote) since no one individual is in a position to know everything about the different components of the interdisciplinary project. In short, we might expect group members to decide what model of collaboration to follow. But even here an individual or set of individuals can occupy authoritative roles. For example, a PI or a set of co-PIs conceive of a project that gets awarded a grant, and then they recruit interdisciplinary experts to work under them. Alternatively, group members may give someone or some sub-set of individuals authority to make decisions with the group's interests in mind to help them function efficiently. For each model, the PI (or someone or sub-group) with authority can make important calls for how cognitive integration will occur. For integration by leader (where one individual scientist is tasked with integrating each collaborator's contributions), the PI can designate who will integrate the knowledge or perhaps they will do it themselves. For common group learning (where all collaborators learn the same content together), the PI can decide what the group needs to learn. For negotiation (where cognitive integration is negotiated iteratively among collaborators with different expertise), the PI may be the final arbiter of

[84] Though one can imagine a case where someone in a junior position would be willing to go along with data fabrication or fraud because of undue pressure from someone in a position of authority.

any outstanding conflicts. Finally, for modelling (where the relevant content is "fed" into a preferred model), the PI can choose (or have the strongest say on) which model to use.

My examples above focus on groups where a PI (or person or sub-group with authority) decides how to balance dependence relations. If the PI determines expertise to be most important, then everyone, including the PI, will concede to the relevant expert's recommendations. Epistemic dependence is thus prioritized over the authority of certain social roles. Or, they may decide that while expertise is important, individuals with more intellectual success (e.g., successful publications, grant proposals, and academic awards), broadly construed, are in a better position compared to those with less intellectual success to help improve papers and projects. So individuals in social roles with more authority, and by extension often more general experience, can intervene over those with the relevant expertise. These examples show how epistemic dependence relations can be altered by individuals in authoritative social roles. This is significant because differential degrees of authority are common in science's collaborative practices.

However, one concession needs to be made about the relationship between these dependence relations: they can lead to conflict among group members. This will depend on the social dynamics of a given collaboration, but conflicts are neither hard to imagine nor are they implausible. Junior collaborators may, for example, disagree with a PI's decisions because they believe the PI does not have the research expertise that, say, a postdoc has. Alternatively, the PI may feel frustrated that their junior collaborators are too often questioning their recommendations, even though they are broadly the most experienced and hold the most authority. There can also be cases where junior collaborators do not receive sufficient credit for their work because they are in a disadvantaged bargaining position for first author (see Bruner and O'Connor 2018).[85] My aim is not to address these conflicts here, but I do want to highlight that in recognizing the interaction between epistemic dependence and social roles, there are also important tensions between them.[86]

My discussion so far may raise some questions. Should we be concerned that PIs can, in virtue of the authority assigned to their role, alter group members' epistemic dependence relations with each other? Shouldn't epistemic dependence relations be organized according to expertise? And why should collaborators concede to the PI, if someone else in the group – e.g., the postdoc – is the expert? These questions raise broad normative concerns about the organization of scientific knowledge production in collaborative groups. Answering them will require considering a group's goals, the composition or organizational structure of members both in terms of expertise and authority, as well as any external constraints (e.g., grant or funding requirements). But it is worthwhile to note that idealistic intuitions about how science should work – e.g., that deference should only be given to experts – are often unrealistic and unproductive in many aspects of scientific practice. Scientists partake in collaborative

[85] Bruner and O'Connor's (2018) modelling results demonstrate how collaborators in authoritative positions have bargaining advantages over junior collaborators.
[86] I will address some of these conflicts in Chapter 5.

efforts because it is more productive and efficient than working alone – it pays off to be collaborative and dependent on others (see Boyer-Kassem et al 2018; Katz and Martin 1997; Lee and Bozeman 2005). Moreover, collaborations often involve people at different stages of their scientific career. It generally makes sense not only to divide cognitive labor, but also to designate authority to some and not others to ensure goals are completed efficiently and in good time. Individuals at more advanced stages of their careers often have more experience navigating broader aspects of research, like how to divide work efficiently and dealing with bureaucratic obligations like grant writing and obtaining IRB approval, to name a few. Just because someone is the content expert, does not mean that they are good at organizing and managing the progress of others. A junior person with the most expertise may not know how to navigate the broad aspects of research mentioned, because these responsibilities tend to fall on more senior researchers as their careers advance. Indeed, it may be hard to get things done if someone without the relevant general professional experiences is put in charge. Idealistic intuitions about how collaborations should work, implicit in the questions raised above, would need further explanatory justification within scientific practice to be seriously considered. My aim here has not been to address normative concerns yet, but to show that epistemic dependence and social roles can and do notably share an interactive relationship, and that they may sometimes conflict.

IV Longino's Norms of Inquiry for Collaboration?

My discussion of epistemic dependence and social roles has focused on demonstrating that they are important components of collaborative practice, and that attention needs to be paid to each as well as to how they interact. We are now in a position to consider norms for well-functioning collaboration, which include both the production of reliable empirical content, as well as positive social relations. Any norms for collaboration would need to consider relations of epistemic dependence, social roles, and their interactions, given their centrality in collaborative practice. Norms that do not consider these features are unlikely to be effective in addressing concerns related to the internal dynamics of such groups. But before any proposals, perhaps there are existing normative accounts of inquiry in the philosophy of science literature that I could use to apply to collaborative groups. One influential feminist normative account of inquiry is Helen Longino's "critical contextual empiricism."[87] This account is intended to apply to interaction across or between different groups or communities, but perhaps its norms can be applied to collaborative groups? I will consider whether Longino's normative criteria for scientific inquiry can be applied to collaborations in a way that accommodates relations of epistemic dependence, social roles and their interactions. My analysis will show, however, that Longino's normative criteria may not be directly applicable because they do not include consideration about authority structures in scientific collaborations. I start with an overview of Longino's account.

[87] Also see Harding (2015); Wylie (2012).

Longino argues that scientific inquiry is more productive when it includes diverse perspectives and approaches. Communities or groups with different background assumptions can help bring problematic or unexamined assumptions – e.g., sexist, racist, and ableist ones – to the forefront through critical interaction. The group receiving critique must also adequately respond – so there ought to be sufficient "uptake" of criticism. Longino writes: "Effective critical interactions transform the subjective into the objective, not by canonizing one subjectivity over others, but by assuring that what is ratified as knowledge has survived criticism from multiple points of view" (2002, 129). If a view survives critique from various perspectives, then it is deemed more objective than a competing view that has not. Prioritizing critical interaction also helps ensure that marginalized perspectives are not overlooked or hastily dismissed. Longino proposes the condition of "tempered equality" of intellectual authority, which requires that different perspectives are given equal opportunity to take part in the critical interactions of a community.[88] Longino writes: "[A]ny perspective has a prima facie capacity to contribute to the critical interactions of a community, though equal standing can be lost owing to failure to engage or to respond to criticism" (2019). Different perspectives are initially given equal consideration, but they will not continue to receive attention if members of that community fail to engage with relevant critiques. Furthermore, such equality must be "qualified or tempered" in terms of intellectual authority or the "cognitive or intellectual skills of observation, synthesis, or analysis, which enable one to make cogent comments about matters concerning which one knows less than another" (2002, 133 fn19). Note that intellectual authority refers to broad cognitive or intellectual skills, rather than narrow expertise. By contrast, "cognitive authority" does refer to expertise or "[…] the amount of knowledge one has, and is thus fairly domain-specific. One can be cognitively authoritative with respect to matters in astrophysics but ignorant of cell biology" (2002, 133). Intellectual authority and cognitive authority can often be considered together but should not be conflated. As Longino puts it: "[T]he condition [tempered equality] speaks of intellectual authority, not cognitive authority. It is compatible with according greater cognitive authority on some matters to those one regards as having acquired more knowledge concerning those matters than others" (2002, 133). Overall, tempered equality is a condition that ensures the equal consideration of various contributions or perspectives, tempered in terms of *intellectual authority*. This helps prevent marginalized perspectives from being ignored because only intellectual authority can be tempered, and not someone's race, gender, and/or class identity.

Critical interaction must also take place in appropriately shared "venues," such as peer-reviewed scientific journals and conferences, which offer public arenas for engagement. These venues help ensure that the relevant "public standards" are followed, such as shared criteria of evidential evaluation. An opinion piece with a controversial conclusion (e.g., climate change is a hoax to promote a globalized

[88] Longino writes: "Thus a community must not only treat its acknowledged members as equally capable of providing persuasive and decisive reasons and must do more than be open to the expression of multiple points of view; it must also take active steps to ensure that alternative points of view are developed enough to be a source of criticism and new perspectives. Not only must potentially dissenting voices not be discounted; they must be cultivated" (2002, 132).

shadow government) funded by corporate interests and published online without peer-review would thus not count. Members of the scientific community are not obligated to engage with content that does not follow shared norms of inquiry. But a well-functioning scientific community does need groups with diverse perspectives for genuine critical interaction. By contrast, a homogenous scientific community with unchallenged perspectives would not be considered objective, whereas one that includes and engages different perspectives would. In sum, Longino's account provides what Heidi Grasswick calls "internal democratic conditions [for] epistemic communities" (2018).

Let's now apply Longino's normative conditions to collaborative groups, focusing on the uptake (of criticism) and tempered equality of intellectual authority conditions.[89] I apply these norms to a case of co-authorship collaboration. This group is comprised of a PI, a postdoc, and a graduate student researcher. The postdoc is designated lead writer, and they have solicited and received comments from their fellow collaborators. Assume in this case the PI has the most expertise, while the graduate student has the least. In receiving everyone's comments, however, the postdoc notices that the PI's and the graduate student's comments conflict significantly in important places. The postdoc adjudicates this discrepancy by being more attentive to the PI's comments. Their rationale is that the PI has more research and publishing experience than the graduate student. In addition, they receive funding from the PIs grant, so they have an obligation to tailor the paper to reflect or at least be complementary with the grant proposal. The postdoc also understands the group's hierarchical structure: the PI has the most authority, while they in turn have more authority than the graduate student. Considering epistemic dependence, the postdoc is more dependent on the PI than the graduate student researcher since the former has more research experience. And considering social roles, the PI's role confers greater authority than the graduate student.

Shifting to the application of Longino's norms, how might the condition of tempered equality, which promotes the equal consideration of epistemic contributions while tempering intellectual authority, work in this case? Since the PI arguably possesses broader skills of "intellectual analysis, observation, and synthesis" compared to the graduate student, it seems the PI's comments should be given more weight. The graduate student, on the other hand, is working in the PI's lab to learn and gain experience. It seems understandable then that the postdoc is more attentive to the PI's comments.[90] Note, however, that the PI's and graduate student's epistemic contributions are not weighted unequally because of their social identity (e.g., gender, race, or sexual orientation). As Grasswick puts it: "[…] objectivity can only be successfully pursued in situations where knowers are respected for their epistemic

[89] The condition of public venues does not seem especially relevant here given that we are considering the internal dynamics of a collaborative group. While the shared standards condition for, say, evidential evaluation in this case can easily be met by appealing to the standards of the research community that this group is intending to produce work in.

[90] Longino clarifies tempered equality by writing: "While the criterion imposes duties of inclusion and attention, it does not require that each individual, no matter what their past record or state of training, should be granted equal authority on every matter" (2002, 132-133).

contributions and not ignored or assigned less credibility because of who they are" (2018). However, epistemic contributions can *justifiably* be weighed less if an individual has less intellectual authority compared to others. Here social roles have an important purpose when assessing this because they can signal or communicate an individual's experience and skillset. PIs, for example, usually have a successful publication track record, and they are awarded grants partly because of this.[91] Equality can therefore be tempered because of who an individual is *in the context of a collaboration*: their assigned social role in the group helps signal their research experience and skillset. It seems then that tempered equality can be met in collaborations where members have differential authority so long as the unequal authority structure mirrors what results from tempering members' intellectual authority.

The other condition, uptake, requires that there be genuine responsiveness between critics and defenders of a particular view. The contents of the espoused view should change over time in response to critical comments (Longino 2002, 129-130).[92] The postdoc, as lead writer, needs to seriously consider and be responsive to both the PI's and graduate student's critical contributions to the paper. A measure of whether the postdoc has adequately given uptake to their colleagues' contributions is if the paper is transformed in light of feedback. The final version of the paper would need to have undergone and survived multiple rounds of criticisms from various perspectives. However, satisfying tempered equality in this example involves ranking collaborators' epistemic contributions based on intellectual authority. How might the postdoc then balance this with the consideration of diverse perspectives required by the uptake condition? This balance is more straightforwardly achieved when collaborators occupy equal social roles with no difference in authority. In democratic groups, the postdoc would consider diverse criticisms from all of their colleagues. Disagreements would be resolved through something like deliberation and negotiation among members. But the co-authorship example involves differential degrees of authority, which makes it hard for the postdoc to give complete uptake to both sets of comments, especially since the PIs and graduate student's comments conflict. The postdoc will be more responsive to the PI's comments than the graduate student, given the PIs greater intellectual authority. While this seems like a reasonable choice on the postdoc's part, it raises questions about the application of tempered equality and uptake in hierarchically structured groups. It is unclear when and how to satisfactorily balance ranking different epistemic contributions based on intellectual authority, while also meeting the democratic aim of giving uptake to a range of critical perspectives.

The co-authorship example reveals a tension with Longino's tempered equality and uptake conditions when applied to collaborative practices involving differential degrees of authority. In cohesive collaborations like co-authorship, critical interaction will occur on a more selective basis, favoring

[91] Postdocs, on the other hand, have completed their doctorate training, and they are on their way to becoming future PIs if they are able to publish respectable work. Finally, graduate students are in the early stages of their academic career, learning from collaborators, including peers and those with more experience like PIs and postdocs.

[92] Such changes "[…] may comprise the acceptance of different beliefs, the modification of beliefs, the development of new data, reasons, and arguments" (Longino 2002, 130).

certain voices over others. This potentially strains a broad aim of Longino's account to promote diverse critical interaction. An important reason why though is that it seems Longino's account was not intended for socially cohesive groups like collaborations. This suggests that the account is perhaps not appropriately applicable when considering such groups. Moreover, it is unclear whether Longino would accept this application. Tempered equality and uptake were intended to promote equality of intellectual authority, and exposure to "the broadest range of criticisms" (2002, 132). They are supposed to help mitigate systemic exclusionary practices that ignore people based on their social identities (e.g., race, gender, ethnicity, and class), while valuing dominant Western and European science as the best (2002, 132). While Longino's account focuses on differing social identities, it says little about differing social roles within the internal practices of science.

Also, note that even though I'm suggesting Longino's two conditions are perhaps not a good fit for collaboration in science, I'm not claiming that because scientists do not abide by Longino's norms when they collaborate the norms are wrong.[93] To be clear, my discussion is not a direct criticism of Longino's account. It is instead an attempt to apply her norms to groups with authority structures. This is a central consideration for my purposes since, for example, the bulk of publications are produced by scientists working in collaborations with sophisticated divisions of cognitive labor, both in terms of epistemic efforts and social roles with differential degrees of authority. The hierarchical structure of collaborative groups which allow for differential weighting of contributions are important. One reason is that collaborations often involve members at different career-stages. This, for example, allows junior collaborators to receive guidance and mentorship from senior collaborators to learn a variety of skills, such as empirical techniques, how to write a publishable paper, and how to identify projects that would receive uptake from working scientists. Senior collaborators, in turn, benefit from the help of junior collaborators by distributing some of their labor to them. When collaborations work well, members learn from one another, produce joint research products, and junior members acquire skills as a result. Second, senior collaborators are often grant holders who provide funding for their junior collaborators, like postdocs and graduate students. Since senior collaborators are answerable and responsible to funding agencies about the contents produced and the group's progress, it seems reasonable that their contributions be given important attention. These are some reasons why we need normative conditions that include consideration about authority structures and allow for the differential weighting of contributions. Such norms also need to aid in mitigating the potential problems that arise when there are differential power dynamics among collaborators. I consider some of these issues in the next chapter.

[93] Thanks to Conor-Mayo Wilson for helping me clarify this point.

Conclusion

In this chapter, I argued that collaboration in science is distinguished by the dependence relations collaborators share with one another – in particular, epistemic dependence and social roles. Furthermore, when attention is paid to how both of these features operate in collaboration, we see that they interact and can sometimes come into conflict. I then considered whether Longino's normative conditions for objective knowledge acquisition could be applied to collaborations. I argued that the conditions of uptake and tempered equality can come into tension when considering groups where members possess differential degrees of authority. We thus need a different set of normative conditions that directly consider the sorts of group dynamics and structures exhibited in collaboration. In the next chapter, I propose some working normative conditions for scientific collaboration, which I then use to assess the IPCC group and finally offer an answer to the skeptics and members of the public for why they should trust the IPCC's collaboratively produced consensus position.

Bibliography

Andersen, Hanne, and Susan Wagenknecht. 2013. “Epistemic Dependence in Interdisciplinary Groups.” *Synthese* 190: 1881–98.

Biagioli, Mario. 2003. “Rights or Rewards? Changing Contexts and Definitions of Scientific Authorship.” In *Scientific Authorship: Credit and Intellectual Property in Science*, edited by Mario Biagioli and Peter Galison, 253–80. New York: Routledge.

Bright, Liam Kofi, Haixin Dang, and Remco Heeson. 2018. “A Role for Judgment Aggregation in Coauthoring.” *Erkenntnis* 83: 231–52.

Bruner, Justin, and Cailin O’Connor. 2018. “Power, Bargaining, and Collaboration.” In *Scientific Collaboration and Collective Knowledge*, edited by Thomas Boyer-Kassem, Conor Mayo-Wilson, and Michael Weisberg, 135–57. New York: Oxford University Press.

CASRAI. 2014. “CRediT - Contributor Roles Taxonomy.” CASRAI. 2014. https://casrai.org.

Cho, Adrian. 2011. “Particle Physicists’ New Extreme Teams.” *Science* 333 (6049): 1564–67.

Collins, Harry, and Robert Evans. 2002. “The Third Wave of Science Studies: Studies of Expertise and Experience.” *Social Studies of Science* 32 (2): 235–96.

Collins, Harry, Robert Evans, and Mike Gorman. 2007. “Trading Zones and Interactional Expertise.” *Studies in History and Philosophy of Science*, no. 38: 657–66.

Fagan, Melinda Bonnie. 2012. “Collective Scientific Knowledge: Collective Scientific Knowledge.” *Philosophy Compass* 7 (12): 821–31.

Galison, Peter. 2003. “The Collective Author.” In *Scientific Authorship: Credit and Intellectual Property in Science*, edited by Mario Biagioli and Peter Galison, 325–55. New York: Routledge.

Grasswick, Heidi. 2018. “Feminist Social Epistemology.” *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/feminist-social-epistemology/#DemStrKnoPro.

Harding, Sandra. 2015. *Objectivity and Diversity: Another Logic of Scientific Research*. London: University of Chicago Press.

Hardwig, John. 1985. “Epistemic Dependence.” *The Journal of Philosophy* 82 (7): 335–49.

———. 1988. “Evidence, Testimony, and the Problem of Individualism - A Response to Schmitt.” *Social Epistemology*, Symposium on Epistemic Interdependence, 2 (4): 309–21.

———. 1991. “The Role of Trust in Knowledge.” *The Journal of Philosophy* 88 (12): 693–708.

Huebner, Bryce, Rebecca Kukla, and Eric Winsberg. 2018. “Making an Author in Radically Collaborative Research.” In *Scientific Collaboration and Collective Knowledge*, edited by Thomas Boyer-Kassem, Conor Mayo-Wilson, and Michael Weisberg, 95–116. New York: Oxford University Press.

ICMJE. 2020. “Defining the Role of Author and Contributions.” International Committee of Medical Journal Editors. 2020. http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html.

Katz, J. Sylvan, and Ben R. Martin. 1997. “What Is Research Collaboration.” *Research Policy* 26: 1–18.

Kitcher, Philip. 1990. “The Division of Cognitive Labor.” *The Journal of Philosophy*, no. 87: 5–22.

Lariviere, Vincent, Chaoqun Ni, Yves Gingras, Blaise Cronin, and Cassidy R. Sugimoto. 2013. “Global Gender Disparities in Science.” *Nature* 504 (7479): 211–13.

Lee, Sooho, and Barry Bozeman. 2005. “The Impact of Research Collaboration in Scientific Productivity.” *Social Studies of Science* 35 (5): 673–702.

Longino, Helen E. 1990. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton: Princeton University Press.

———. 2002. *The Fate of Knowledge*. New Jersey: Princeton University Press.

———. 2019. “The Social Dimensions of Scientific Knowledge.” *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/scientific-knowledge-social/.

Nature Research. 2019. “Authorship.” Nature. 2019. https://www.nature.com/authors/policies/authorship.html.

Rossini, Frederick, and Alan Porter. 1979. “Frameworks for Integrating Interdisciplinary Research.” *Research Policy* 8: 70–79.

Science. 2020. “Science Journals: Editorial Policies.” CASRAI. 2020. https://www.sciencemag.org/authors/science-journals-editorial-policies.

Shapiro, Scott J. 2014. “Massively Shared Agency.” In *Rational and Social Agency: The Philosophy of Michael Bratman*, edited by Manuel Vargas and Gideon Yaffe, 257–93. New York: Oxford University Press.
Staley, Kent. 2007. “Evidential Collaborations: Epistemic and Pragmatic Considerations in ‘Group Belief.’” *Social Epistemology* 21 (3): 15.
Tuomela, Raimo. 1992. “Group Beliefs.” *Synthese* 91 (3): 285–318.
Urfalino, Phillipe. 2006. “Apparent Consensus and Voting: Modes of Collective Decision-Making.” https:// www.academia.edu/4277481/Apparent_Consensus_and_Vote.
Venkatraman, Vijaysree. 2010. “Conventions of Scientific Authorship.” *Science Magazine*, April 16, 2010. https://www.sciencemag.org/careers/2010/04/conventions-scientific-authorship.
Wagenknecht, Susan. 2015. “Facing the Incompleteness of Trust: Managing Dependence in Scientific Practice.” *Social Epistemology* 29 (2): 160–84.
Weisberg, Michael. 2009. “New Approaches to the Division of Cognitive Labor.” In *New Waves in Philosophy of Science*, edited by P.D. Magnus and Jacob Busch, 250–69. New York: Palgrave Macmillan.
Weisberg, Michael, and Ryan Muldoon. 2009. “Epistemic Landscapes and the Division of Cognitive Labor.” *Philosophy of Science* 76 (2): 225–52.
Wray, K. Brad. 2007. “Who Has Scientific Knowledge.” *Social Epistemology* 21 (3): 337–47.
———. 2018. “The Impact of Collaboration on the Epistemic Cultures of Science.” In *Scientific Collaboration and Collective Knowledge*, edited by Thomas Boyer-Kassem, Conor Mayo-Wilson, and Michael Weisberg, 117–34. New York: Oxford University Press.
Wylie, Alison. 2012. “Feminist Philosophy of Science: Standpoint Matters.” *Proceedings and Addresses of the APA* 86 (2): 46–76.
Zollman, Kevin. 2010. “The Epistemic Benefit of Transient Diversity.” *Erkenntnis* 72: 17–35.

5 Normative Criteria for Collaboration

I argued previously that dependence relations – epistemic dependence and social roles – are important aspects of contemporary collaborative practices in science. Dependence relations are interactive, and they sometimes can be in tension. In particular, the structural organization of collaborative groups can vary depending on whether expertise or authority of particular social roles is given priority. I also considered Helen Longino's normative conditions for scientific inquiry as a potential framework to assess scientific collaboration, given their impact and wide-ranging appeal among philosophers. However, the analysis does not yield a successful application because the account does not include consideration of authority structures exhibited in collaborative groups. Therefore, my aim in this chapter is to offer some normative conditions for well-functioning collaboration that consider the social dynamics and organizational structures of scientific collaborations.

Recently, some philosophers have raised and discussed concerns about the groupthink phenomenon in socially cohesive and highly interactive groups like collaborative teams in science (Solomon 2006; Tollefsen 2006). I use this discussion as a starting point to consider normative conditions for well-functioning collaborative groups that aim to mitigate negative social dynamics. An important consequence of the conditions proposed here is that collaborative groups should not rely on democratic standards of equality as a guiding norm. The structural orientation of hierarchically organized groups makes them inherently at odds with ideals that promote equality among group members. I also assess the IPCC collaboration using these norms and argue that the group's procedures and practices generally satisfy them. This then puts us in a position to answer the climate change skeptic, in a way importantly different than the reasoning provided by Oreskes, regarding why we should trust the IPCC's consensus position on anthropogenic climate change.

This chapter is organized into four sections. In section I, I discuss concerns about negative social dynamics in scientific collaboration like groupthink and abuse of power. In section II, I propose four normative criteria for well-functioning collaboration, which take into consideration the vulnerabilities of socially cohesive groups. Section III then uses the proposed criteria to assess the IPCC's collaborative efforts. I show that the IPCC has procedures and practices that generally satisfy the normative criteria. I also identify some places for improvement. Finally, in section IV, based on my analysis of the IPCC collaboration, I offer a positive reason for why their collaboratively produced consensus position should be trusted.

I Mitigating Negative Social Dynamics

In the last chapter, I argued that in collaborations where members have differential degrees of authority, Longino's conditions of tempered equality and uptake of criticism can come into tension. Tempered equality allows one to rank members' contributions based on intellectual authority, while uptake (of criticism) requires the equal consideration of critical feedback. However, this analysis has implications for some collaborative consensus models too, such as John Beatty and Alfred Moore's (2010) "deliberative acceptance." [94] Beatty and Moore's model shares a similarity with Longino's uptake condition. It also emphasizes the importance of equality among group members for effective deliberation to take place. Beatty and Moore highlight the centrality of equal standing by adding that a vote can function as the deciding factor when the group cannot reach deliberative acceptance. This model also does not authority structures in hierarchically organized collaborations. It would be unlikely for a PI, whose team consists only of junior collaborators, to ask their team to cast a vote, with the assumption that the undergraduate researcher's perspective is equal to their own, the postdoc's, and graduate student's. So why does this matter? Authority structures should be a central consideration for potential normative criteria for scientific collaboration. The criteria I propose, contrary to the account previously discussed, do not require the equal standing of group members. Rather, I start from the assumption, consistent with important collaborative practices in science, that there is often an inequality of power among group members occupying different social roles. In short, these normative criteria will be sensitive to the relational and hierarchical dynamics of scientific collaborations.

In considering norms for collaboration, a useful place to start is with the sorts of social dynamics that hinder productive working relations for these groups. For example, dynamics that prevent consensus from being achieved and dynamics that facilitate the appearance of consensus without there being a genuine consensus of the individuals involved. I will focus on the latter. Groups that engage in collaboration tend be to "socially cohesive" in the sense that members can be highly interdependent (Tollefsen 2005, 41).[95] This, for example, is exhibited in the ways in which collaborators can be relationally dependent on one another whether epistemic, authority-related, or both. They also operate as a unit and they share joint goals and commitments. Socially cohesive groups, however, can be vulnerable to phenomenon like groupthink, which some philosophers have discussed especially where deliberation is involved (Tollefsen 2006; Solomon 2006; Wylie 2006). Groupthink (Janis 1972) occurs when group members conform to a view because of peer pressure and/or coercion from more authoritative members.[96] It can notably lead to the suppression of dissent. During collaboration, groupthink can occur, for example, when the group deliberates about their joint consensus position or when they decide

[94] See Chapter 1 for a discussion of Beatty and Moore's account.
[95] As Tollefsen writes: "To the extent that a group maintains stability of agency by forming shared values and being mutually responsive to the intentions and commitments of the participants, it will be more socially cohesive than those groups whose agency is less interdependent" (2006, 42).
[96] Also see Chapter 1 for a brief discussion of groupthink.

authorship order for a publication. Absent good policies or ethical leaders, it seems an easy possibility for some members to feel pressure from certain individuals to consent to actions that they otherwise would not choose.

In addition, hierarchical collaborations can make it easy for individuals in powerful roles to abuse their authority with impunity. Credit attribution is one place where this can occur. Authoritative individuals can profit from others' intellectual work, for example, by taking credit for work they did not do. And worse, junior collaborators can feel beholden to the whims of these authoritative individuals when their future prospects depend heavily on these individuals' professional connections and letters of recommendation. When considering what counts as well-functioning collaboration and what factors to mitigate, phenomena like groupthink and abuse of authority should be at the forefront. Hierarchies of social roles are not bad simpliciter, but without appropriate mitigating norms it can leave members with little or no authority vulnerable to wrongful or unjust actions from senior members and others who have strong allegiance to them. This can disrupt and erode positive group dynamics like mutual trust, mutual responsiveness, and productivity.

II Four Normative Criteria

In this section, I propose four normative criteria that are sensitive and specific to the social dynamics and practices of scientific collaboration. They each highlight an important aspect of collaboration and offer some ways to mitigate problems that may arise. "Fair credit attribution" addresses concerns related to the distribution of credit. "Respectful participation" proposes ways for group members to be treated with respect, while also acknowledging the inequality of authority among members. "Group leader accountability" holds group leaders or those in positions of authority accountable for their actions as a measure to mitigate abuse of power. And "external grievance reporting" offers an avenue for group members, especially junior and members in vulnerable positions, to report concerns or problems to an independent body. These criteria are attuned to protecting the most vulnerable collaborators, such as junior scholars, and minority members (e.g., people of color, women, and individuals with underprivileged backgrounds).

I also want to emphasize that these normative criteria are intended for collaboration in science, so groups must adhere to other norms for scientific inquiry connected to the shared empirical and evidential standards of the scientific disciplines or fields in which the work is situated.[97] Without this constraint, collaborative groups that, for example, promote conspiracy theories about climate change could in theory satisfy the four criteria. But they should not satisfy the criteria for well-functioning collaboration precisely because their claims do not adhere to the empirical and evidential standards of climate science research. Indeed, many of these purported groups of dissenting scientists neither publish

[97] Thanks to Conor Mayo-Wilson for this suggestion.

nor cite peer-reviewed science. Here, Longino's articulation of the public standards condition is especially helpful.[98] The condition is as follows: "There must be publicly recognized standards of reference to which theories, hypotheses, and observational practices are evaluated and by appeal to which criticism is made relevant to the goals of the inquiring community" (Longino 2002, 130).[99] Having something like a public standards requirement upfront helps ensure collaborative groups comprised of scientists, whether making a public policy recommendation or publishing a paper, meet the shared empirical and evidential standards of their relevant research communities.

In addition, one caveat is that I expect to clarify and elaborate these criteria over time. I have focused on particular cases and practices of collaboration that result in consensus such as the IPCC's consensus position, co-authorship practices, and research collaboration in neural engineering. The relevant point of similarity among these different cases for my purposes is the power hierarchies that exist among group members.[100] However, I recognize that there are other potential axes of analysis, such as variation with respect to authorship practices. More broadly, I anticipate that growing awareness of the practices of other "disciplinary ecologies" will result in amendments to these normative criteria for fruitful application.[101] Let's start with the first condition concerning credit.

1. ***Fair Credit Attribution***. Distribution of labor among collaborators should be accurately reflected when crediting collaborators for their work, especially when listing co-authors, contributors, and their contributions. This condition helps promote well-functioning collaborations and deters author disputes, especially in cases where junior collaborators are shortchanged for their work. It emphasizes the importance of responsible conduct in one's role as both as a collaborator and co-author.[102] For example, authors should be familiar with the most current best practices on shared authorship, such as the guidelines offered by COPE, the Contributor Roles Taxonomy (CRediT), and the International Committee of Medical Journal Editors' (ICMJE) definitions for who count as authors and contributors.[103] This could be a standard component of graduate school education with respect to professionalization. COPE makes three recommendations: (a) that a culture of ethical authorship be encouraged and practiced; (b) that authors ought to discuss the topic of authorship regarding their co-authored paper early and throughout the collaboration; and (c) that authors collectively decide on authorship before working on the article (Albert and Wager 2003). Complementary with COPE's recommendations, the Contributor Roles Taxonomy or the CRediT model offers a fine-grained taxonomy of the various contributions researchers make to collaborative projects resulting in publications. The list identifies 14 contributor roles: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation,

[98] Thanks to Carole Lee for this suggestion.
[99] Longino also states that on this condition "individuals and communities adopt criteria of adequacy by which they may be nonarbitrarily evaluated" (Longino 2002, 130).
[100] Thanks to Carole Lee for helping me make this point explicit.
[101] I am borrowing the phrase "disciplinary ecologies" from Mario Biagioli (2003) here.
[102] See Andersen (2014) for a discussion on four cases of misconduct and co-author responsibility.
[103] Thanks to Carole Lee for letting me know about these authorship guidelines.

Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, and Writing – review & editing.[104] This taxonomy, now adopted by a number of leading science journals, offers ways to specify more clearly the work of contributors. Contributors' work can be recognized either in the author line or the acknowledgments section. It is recommended that groups take on the "shared responsibility" of (a) knowing the obligations associated with their role; and (b) taking time to understand and accept their assigned roles (CASRAI CRediT 2014). Finally, the ICMJE recommends four criteria for authorship: (i) significant contribution to the "conception or design" or "acquisition, analysis, or interpretation" of the work; (ii) drafting or revising of the work for "important intellectual content"; (iii) approval of the final version; and (iv) consent to be "accountable for all aspects of the work" (ICMJE 2020). In addition, non-author contributors who do not meet these criteria should be acknowledged for their contributions (e.g., proofreading or editing, general supervision).

These guidelines offered by COPE, the CRediT model, and the ICMJE offer ways to help deter common authorship disputes and, importantly, protects junior authors who are often in a disadvantaged position. For example, following COPE's suggestion of an ethical culture of authorship and/or the ICMJE's authorship criteria would not allow for the dishonest practice of negotiating author credit for either short or long-term benefits, which is sometimes called "gift authorship." This involves cases where people are listed as authors even though they did little to no work, especially to "curry favor" with senior researchers (Albert and Wager 2003, 34). This can also occur in exchange with others to "swell publication lists" (2003, 34). Junior collaborators are especially vulnerable in these circumstances if they have to negotiate with senior collaborators for author credit, given the unequal power imbalance (see Bruner and O'Connor 2018).[105] Of course, power imbalances can also exist between collaborators who are roughly at the same career stage through factors like prestige bias, gender, race, and class dynamics all of which can create unfair advantages for some and disadvantages for others. The guidelines also help deter another type of authorship dispute called "ghost authorship," where people are intentionally left off of the author list even though they qualify as an author (Albert and Wager 2003). For example, COPE cites a case where "the writer of a review article found her name replaced with that of her boss, because she was on maternity leave when the final version was submitted" to the publication venue (2003, 32). However, note that ghost authorship occurs more commonly when industry is involved, such as pharmaceutical companies. In these cases, a company deliberately conceals the name of someone who contributed significantly to an article with a clear aim to benefit the company. Hiding their contribution gives the perception that the article independently supports the company's work (Panter 2020). Authorship disputes like gift and ghost authorship are more common than people may think. At a recent

[104] For a summary of each of these positions see CRediT (2014).
[105] Also see Chapter 4.

COPE panel on responsible authorship, it was noted that gift and ghost authorship make up roughly 19% of authorship conflict cases (Pierson 2019).[106]

Overall, awareness about ethical authorship and a commitment to attributing credit appropriately can mitigate unethical co-authorship practices and help eliminate unfair bargaining contexts. Note, however, that different types of collaboration may call for slightly different practices, such as more granular or broad descriptions of author and contributor roles. My discussion of COPE's recommendations, the CRediT model, and the ICMJE's guidelines aims to showcase some good practices of fairly crediting authors and contributors for their cognitive labor. Finally, I also want to note that while my discussion of fair credit attribution has focused mostly on co-authorship, it is also broadly applicable to collaboration. So even if a collaboration does not have any co-authored papers, members should still be fairly credited for their research contributions, when, for example, content is being discussed whether at conferences or even on a group's website.

2. ***Respectful Participation***. We need to reconceive participation among group members in a way that does not require everyone to be "deliberative equals," as suggested by Beatty and Moore's model, while maintaining respectful interaction. How might we *productively* account for group members' inequality of power or authority due to different social roles and, to use Longino's term, "intellectual authority"?[107] There are two ways in which institutional training or education could be implemented to fulfill this norm. First, place important emphasis on senior collaborators' leadership obligations. Senior collaborators should promote respectful interaction among group members, especially during group deliberation and discussion. This can be done through explicit encouragement of such activity by publicly endorsing the norm that disrespectful behavior towards others will not be tolerated. This can also be done by modelling acceptable behavior when interacting with others. When senior collaborators, for example, routinely belittle or shut down junior members' concerns out of irritation, they are modelling that it is an acceptable way to treat one's collaborators. On the contrary, they can reasonably and respectfully disagree, and offer an explanation. In addition, senior collaborators should devote time and/or space for collective deliberation and discussion for members to share their ideas. This does not mean all voices are equal and that any member can derail plans through dissent. Rather, *participation* from collaborators is taken to be an important component of productive collaboration. It is also important to note that such participation need not be critical in nature; for example, asking a clarifying question could help dispel confusion and be just as productive as a critical comment. So while hierarchically organized groups allow senior collaborators to make executive decisions, which sometimes requires junior members to defer even if they disagree, this should not amount to a culture of fear mongering

[106] Other activities in the list of authorship conflicts include changes to author list (removal, omission, or adding on after submission) at 27%; submission without knowledge 19%; claim of authorship 21%; author order dispute 7%; and forged signatures 7% (Pierson 2019).
[107] Recall that Longino describes intellectual authority as "broad cognitive or intellectual skills" (2002, 133).

where group members are subject to the whims of powerful seniors. Group leaders assume their position precisely because they are expected to (or should) have the competency to make executive judgments on behalf of the group with an aim of benefitting the collaboration, rather than just themselves. Modeling good behavior is key because it sets the tone for what kind of collaborative culture senior collaborators encourage.

Second, I propose emphasizing the normative dimension of social roles. In particular, the obligations that collaborators can reasonably expect of each other in virtue of the social role they occupy in the group. For example, it is reasonable for a group of collaborators to expect the lead writer to be in charge of overseeing a working manuscript, but they cannot expect this from a new undergraduate researcher who just joined the group. We can promote respectful participation by ensuring that group members are given control over the domain of obligations connected to their social role. On this view, one of the ways in which collaborators can participate most fully is when they are given the authority to carry out the obligations expected of them. This sort of freedom over one's assigned domain of work signifies trust and respect in their competence and ability to fulfill their obligations. Respectful participation is not met when a collaborator's competency, on the contrary, is constantly under evaluation and undercut by the doubts of others.[108]

3. ***Group Leader Accountability***. Group leaders should be evaluated and held accountable for their leadership skills by an external body, such as funding agencies and/or the chair(s) of similar or adjacent departments. Evaluators can, for example, check that group leaders do not take credit for their collaborators' work and do not violate fair work climate conditions, for example, by overworking or harassing others. Importantly, a group leader's (perceived) ability to act with impunity is mitigated through regular assessment from other authoritative individuals or bodies.

This recommendation has two important implications. One is the recognition that doing good research is only a necessary, and not sufficient, measure for managing and overseeing the work of others. Senior researchers with a bad track record of managing others and who are not especially good collaborators should not be awarded leadership roles. One possibility for tracking this is for universities or funding institutions to manage personnel files with complaints or grievances. Perhaps this database could be used by the relevant institutions to check an individual's record before awarding them grants that cover postdocs and graduate students. Alternatively, grant awards could be given to groups or teams with at least two individuals designated to oversee the research production of a given project. This would disperse authority across individuals.[109]

The other consequence is the recognition that group leaders bear greater responsibility *for the group* than other members. Two cases of scientific misconduct discussed by Hanne Andersen (2014) will

[108] Note that respectful participation will be somewhat looser if a collaborator is relatively inexperienced. In such cases, junior collaborators should be offered more mentorship by senior collaborators to help with skill-building and to receive feedback.
[109] Thanks to Carole Lee for this suggestion.

help illustrate the point. One example from Germany involves two post-docs who manipulated data for four different publications for which they were listed as first authors. The senior collaborator, Silvia Bulfone-Paus, who was the head of the Department of Immunology and Cell Biology at the Research Center Borstel, reported the allegations and denied any knowledge of misconduct. The investigating committee did not find the senior collaborator guilty of fraudulent activity, but they "concluded that although she had not been involved in the actual data manipulation, she had been grossly negligent as a supervisor to the extent that it constituted scientific misconduct" (2014, 915). Interestingly, this senior collaborator's colleagues wrote an open letter criticizing this verdict, since on their view she was "a victim rather than a culprit" (2014, 915). However, the investigating committee ruled that the first, corresponding, and senior authors were equally accountable for a finished publication, which was in line with publication norms. Moreover, the senior collaborator received "repeated warnings" that questioned the data but was reluctant to investigate. This verdict demonstrates that even though the senior collaborator took no part in the manipulation of data, she nevertheless can be held responsible for the group's published products which she authorized. In this case, her responsibility lies in the fact that she is a senior collaborator and a central co-author who can make authoritative decisions.

Another example involves the head of a social psychology group in the Netherlands, Diederik Stapel, who published many papers in well-respected journals. In 2011, he was found guilty of misconduct and suspended from his university for fabricating experiments and data collection for 85 publications, in part when three of his collaborators reported suspicion about the data. But in contrast to the previous example, the university committees investigating found none of the collaborators to be at fault because: most of them "were graduate students and post-docs, [and] the committees argued that they had to trust the integrity of their supervisor and that, consequently, they could not be expected to take a more critical stance" (2014, 915). This verdict demonstrates that junior collaborators, especially those with little or no authority, do not bear responsibility for the group in the same way as the senior collaborator. In both of these cases, levels of accountability for misconduct are high for senior collaborators who oversee the group, whether it be negligence or manipulation of data.

Overall, this condition highlights the importance of collaborators in authoritative positions, which requires that such authority undergo evaluation from others to prevent misuse or even negligence. Group leaders have a responsibility to encourage and maintain a well-functioning group, which includes the respectful participation of members, and the prevention of misconduct or unethical behavior. Taking on an authoritative role also requires taking responsibility for the group's culture, and the actions of individual members that affect the group. The responsibilities for a group leader are thus substantial and onerous.

4. ***External Grievance Reporting.*** Finally, collaborators, especially junior collaborators, should have a designated and independent avenue to report grievances, such as undue abuse of power, coercion, or misconduct. Reports should be handled by an independent evaluator or team of evaluators with capability to hold those who violate respectful and ethical conduct accountable, including group leaders.

Similar to the previous condition, evaluators might be designated individuals or teams from the funding agency from which the grant for the collaboration was secured, or perhaps someone from the reporting collaborators' university, such as the chair(s) of another department or a designated ombudsperson.[110] This condition provides an external avenue of reporting beyond the group to protect collaborators, especially junior members, from retaliation. It is also a way to maintain the reporting collaborator's trust that the report will be taken seriously and investigated without unfair favoring of someone in an authoritative role. In cases of deep disagreement, the independent evaluator(s) could also serve as a liaison to help resolve conflicts among collaborators. Recall the misconduct case above of Diederik Stapel who manipulated data for many publications, affecting graduate students and post-docs who collaborated with him. Though the report by three of Stapel's junior collaborators revealed his misconduct, Tillburg University's interim report also revealed that other junior faculty and researchers had raised concerns prior to this but no actions were taken (Verfaellie and McGwin 2011). This delay meant that graduate students and collaborators continued to work with fabricated data whether in their own dissertations or as a co-author of Stapel's. The misconduct persisted for a variety of reasons, one of which was Stapel's use of "his power and prestige" (Verfaellie and McGwin 2011). In particular, he was able fabricate data because he alone controlled the entire data collection process and did not let any student or collaborator access it until it was a coded data set. He was "described as intimidating and not tolerating questions from students regarding his refusal to have them involved in data collection" (Verfaellie and McGwin 2011). Another factor was the lack of infrastructure within the university's psychology department to handle reports of scientific misconduct, such as a designated ombudsperson.[111] These factors that helped the misconduct persist suggest the need for a publicly designated grievance reporting process with evaluators who can take concerns seriously and conduct investigations without being affected by factors like the alleged person's influence and prestige.

In sum, the four normative criteria for collaboration I have outlined focus on specific aspects of collaboration and offer ways to mitigate or address problems that can arise. Fair credit attribution promotes appropriate acknowledgement of collaborators' contributions. Respectful participation promotes productive participation (i) by requiring senior members to promote respectful interaction among group members and to model good behavior; and (ii) by encouraging group members respect each other's control over their domain of obligations. Group leader accountability aims to ensure that members with power exercise their authority appropriately and ethically through regular assessments of their leadership skills. Finally, external grievance reporting provides members with an independent

[110] It should be acknowledged the that implementation of this norm will be difficult, and the possibility space sketched here will require further research. Of central importance though is that the reported grievances are handled independently and externally to the parties involved to maintain a level of independence, especially to protect junior collaborators.

[111] This is especially interesting because by-and-large it is universities that handle cases of scientific misconduct (see Retraction Watch: https://retractionwatch.com).

forum to report problems within the collaboration, especially when there are negative or toxic social dynamics connected to inequalities of power or authority.

The proposed normative criteria help mitigate and manage concerns related to negative social dynamics like groupthink and abuse of power in collaborations, especially hierarchically organized ones. First, coercion from authoritative members is mitigated with the group leader accountability, external grievance reporting, and fair credit attribution conditions. Group leader accountability mandates leaders to be answerable to an external body who evaluates their leadership skills, so they are not able to act with impunity. The external grievance reporting condition provides other members, especially juniors, an independent and authoritative avenue to report misconduct or unethical behavior, without feeling vulnerable to retaliation or punishment. And fair credit attribution helps mitigate circumstances related to intellectual contributions and co-authorship credit by not allowing such credit to be either used as bargaining leverage or taken outright by others. Second, peer pressure from other collaborators who are not the group leader can be addressed with respectful participation and external grievance reporting. The former condition requires group leaders to model and maintain a positive work environment for all, which precludes peer pressure to conform since it violates respectful participation. It also encourages group leaders to designate space for group discussion and deliberation with an emphasis on participation (e.g., asking a clarification question or making a suggestion), rather than only making critical comments. The group leader thus has a responsibility to maintain and promote a positive work culture for the group, which includes how other collaborators treat each other. A report can also be filed using the external grievance reporting condition if a collaborator feels that the group leader has not addressed or resolved their concerns. And third, connected to the last two points, if coercion from authoritative members and peer pressure to conform are mitigated, then the suppression of dissent is less likely to take place. Group leaders and collaborators ought not to behave in ways that would create negative group dynamics, such as disrespecting fellow group members. The suppression of dissent is a toxic social dynamic not encouraged by the normative criteria, especially the respectful participation condition.

I should also note that well-functioning collaborations in science are not just those devoid of negative social dynamics. As I mentioned earlier, the group's work must also adhere to the relevant empirical and evidential standards of the research disciplines that the work covers. And in addition other typically positive features that motivate researchers to collaborate should be present. Here is one example: a productive working relationship where collaborators learn from one another and articulate content together in way that they could not have done alone. Another is the successful output of joint products, such as publications, completed research projects or experiments, and conference presentations. Collaborators also build bonds of trust with one another. This could mean, for instance, that in future contexts they will seek the advice of their fellow collaborators as a result of their previous working relationship. Finally, collaborations intermixed with members at different career stages encourages educational exchange. Senior members can vouch for junior members in hiring contexts based on their prior collaboration. Overall, the social cohesion among collaborators, often created by a

commitment to take on a joint task together, has many positive benefits, some of which I've mentioned. But, of course, the intricate social dynamics within collaborations can also make them vulnerable to unproductive social dynamics like abuse of authority and the silencing of marginalized voices. So while collaboration in science ought to be celebrated for the benefits they bring, we must also be cognizant of their problems and work to allay them. The normative conditions outlined above aim to address some of these issues.

Finally, the proposed normative conditions could benefit from further development in two respects. One is that they should be subject to empirical evaluation for effectiveness. For example, qualitative analyses – e.g., a sample of semi-structured interviews, roundtable discussions – of collaborators' experiences working with others to check that the normative conditions accurately capture the social dynamics in collaborations. Policy makers for science, funding institutions, as well as individual researchers and scientists at various career stages will need to take part in this conversation. Effective norms should be constructed in conversation with the relevant communities being affected, and then modified in light of relevant stakeholders' feedback. The other is that these normative conditions, even if they address the right issues, would need buy-in from the relevant scientific institutions and from practitioners of science for them to be effective. They would be utterly useless if treated as another check-list hurdle (sometimes like IRB approval) and would function as window dressing to give the perception that groupthink did not occur. It can be hard to determine whether genuine buy-in from the community has occurred, but a good place to start would be to conduct conversations about these norms with them.

In sum, I've offered four normative criteria for collaboration that take as a basic assumption the inequality of authority among group members, based on factors like assigned social roles and intellectual authority. These norms aim to promote well-functioning social dynamics for hierarchically structured collaborations, while mitigating factors that can lead to toxic social dynamics. More broadly, if a collaborative group promotes and satisfies these norms (in addition to satisfying the relevant empirical and evidential standards), then there are good reasons to consider the group's consensus position trustworthy because it highlights the strengths of the group's social practices and procedures. Even though these conditions will likely be revised depending on the sorts of collaborative contexts and cultures they are applied to, they are an important contribution to the literature because few philosophers have attempted to offer normative conditions for groups without centralizing the idea that members need to be treated as equals for the group to function well, a requirement I have argued is inappropriate in the context of many, if not most, scientific collaborations. In the next section, I will apply these norms to the IPCC's collaboration, and explain why we should trust their consensus position on anthropogenic climate change.

III Assessing the IPCC Collaboration

In Chapter 3, I argued that the skeptics' strategy to undermine the IPCC's consensus position on anthropogenic climate change fails because they invoke the unanimity criterion to assess the consensus position of a collaborative effort. The IPCC's conclusion is not issued from a straightforward voting procedure but from a sophisticated division of cognitive labor involving discussion, deliberation, and negotiation. Further, the IPCC's three Working Groups resemble collaborative groups in scientific practice, and the IPCC writ-large may even be viewed as a large-scale collaboration. However, I still need to show whether the IPCC's collaborative practices meet the conditions of well-functioning collaboration, one that did not emerge as a result of coercive social dynamics. In this section, I will show that they do. I argue that when we look at the group's norms and practices for collaboration that helped lead to the production of their report and their consensus position on anthropogenic climate change, we learn that they generally meet the normative conditions for well-functioning collaboration that I have outlined. I also make some recommendations for improvement. My strategy for defending the legitimacy and trustworthiness of the IPCC's consensus position through an assessment of their norms and practices for collaboration is different from Oreskes's (2004) and Cook et al's (2013) strategy, which focuses on further corroborating the consensus position by aggregating peer reviewed climate science articles.

First, recall that the IPCC is structurally organized into three main Working Groups (I, II, III), as well as a task force, and a task group, in addition to other bodies and individual roles (see Figure 3).

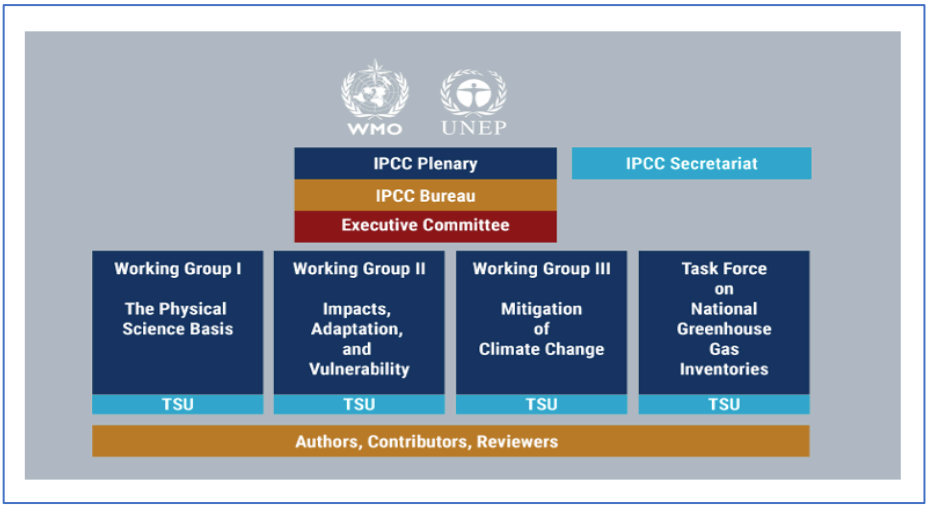

Figure 3: The structure of the IPCC (IPCC 2020b).

Each Working Group produces a report focused on a particular area relevant to members' expertise. For example, Working Group I reports on the current state of the physical components of climate, which include temperature changes in the air, land, and ocean, as well as changes in rainfall, sea level, and glaciers and ice sheets. There are also some unique features about IPCC collaboration worth highlighting.

First, the IPCC's norms for authorship are somewhat different from authorship practices intended for peer review publications in research journals. The group's report undergoes an extensive peer review process (arguably even more extensive than regular peer review processes for journal publication (see below)) but it is intended for a wider audience, which include governmental bodies and organizations involved in public policy decision-making, as well as the public. The IPCC itself also does not conduct any original research, though it's authors, contributors, and reviewers do in their capacity as climate science researchers. In addition, the IPCC's authors, contributors, and reviewers volunteer their time to help report on the most current science and information about the status of anthropogenic climate change. With these features in mind, let me now apply the proposed norms for collaboration to the IPCC group, starting with fair credit attribution.

Fair credit attribution minimizes authorship disputes by requiring that cognitive labor and credit be accurately reflected, such as in the authorship line or the contributions and acknowledgements sections. Responsible conduct is of central importance, and collaborators should neither be bargaining for credit nor shortchanged on credit for work they completed. Collaborators should thus have a clear sense of their responsibilities as author or contributor. So what do the IPCC's authorship practices look like? The selection of authors among a pool of volunteers begins with a call for nominations and submissions of CVs to governments and related organizations (IPCC 2013b, 1). Author selection is based on expertise, which include "scientific, technical, and socio-economic views and backgrounds" (2013b, 1). Each IPCC Working Group divides sub-groups into Chapter teams, which include Lead Authors, Coordinating Lead Authors, and Review Editors (see Figure 4).

(a) **Lead Authors** are responsible for synthesizing material from the most current empirical literature, then producing or writing the report based "the best scientific, technical and socio-economic information available" (IPCC 2013a, 14). Lead Authors work together in small teams to produce and coordinate efforts efficiently. Note that given the demanding nature of this particular role, the IPCC makes a point of emphasizing that such work will be given appropriate recognition. They write:

> The task of Lead Authors is a demanding one and in recognition of this the names of Lead Authors will appear prominently in the final Report. During the final stages of Report preparation, when the workload is often particularly heavy and when Lead Authors are heavily dependent upon each other to read and edit material, and to agree to changes promptly, it is essential that the work should be accorded the highest priority. (2013a, 14)

(b) **Coordinating Lead Authors** share the same responsibilities as Lead Authors, but they are also required to coordinate and ensure consistency across major sections of the report. The names of Coordinating Lead Authors are also acknowledged in the Report.

(c) **Review Editors** of IPCC Reports help the Working Group/Task Force Bureaux to identify appropriate expert reviewers for the report, ensure reviewer comments are taken into consideration and addressed, and provide Lead Authors with advice on handling "contentious/controversial issues and

ensure genuine controversies are reflected adequately" (2013a, 15). Review Editors are also acknowledged in the Report.

Chapter 1: Framing, context, methods

	Last Name	First Name	Role	Gender	Country	Citizenship	Current Affiliation
1	CHEN	Deliang	CLA	M	Sweden	Sweden	University of Gothenburg
2	ROJAS	Maisa	CLA	F	Chile	Chile	University of Chile
3	SAMSET	Bjørn H.	CLA	M	Norway	Norway	CICERO Center for international climate research
4	COBB	Kim	LA	F	USA	USA	Georgia Inst. of Technology
5	DIONGUE-NIANG	Aïda	LA	F	Senegal	Senegal	Agence Nationale de l'Aviation Civile et de la Météorologie (ANACIM)
6	EDWARDS	Paul	LA	M	USA	USA	Stanford University and University of Michigan
7	EMORI	Seita	LA	M	Japan	Japan	National Institute for Environmental Studies, Japan
8	FARIA	Sergio Henrique	LA	M	Spain	Brazil	BC3 - Basque Centre for Climate Change
9	HAWKINS	Edward	LA	M	UK	UK	National Centre for Atmospheric Science, University of Reading
10	HOPE	Pandora	LA	F	Australia	Australia	Bureau of Meteorology
11	HUYBRECHTS	Philippe	LA	M	Belgium	Belgium	Earth System Science & Departement Geografie, Vrije Universiteit Brussel
12	MEINSHAUSEN	Malte	LA	M	Australia	Germany	The University of Melbourne
13	MUSTAFA	Sawsan	LA	F	Sudan	Sudan	Ministry of Animal Resources /Range and Pasture General Directorate
14	PLATTNER	Gian-kasper	LA	M	Switzerland	Switzerland	Swiss Federal Research Institute for Forest, Snow and Landscape WSL
15	TREGUIER	Anne Marie	LA	F	France	France	CNRS
16	CHUERSUWAN	Nares	RE	M	Thailand	Thailand	Suranaree University of Technology
17	HEGERL	Gabriele C.	RE	F	UK	Germany	Geosciences

Figure 4: Chapter 1 team for Working Group 1 of the IPCC's sixth assessment report (IPCC 2020c).

In addition, there are Contributing Authors, Expert Reviewers, and Government Focal Points who make contributions to the Report.

(d) **Contributing Authors** are responsible for preparing technical information through "text, graph, or data," which the Lead Author then incorporates into the Report. Contributing Lead Authors and Lead Authors seek out the help of Contributing Authors for their specialized knowledge. There can be up to many hundreds of Contributing Authors. Their work is also acknowledged in the Report.

(e) **Expert Reviewers** provide comments on the content of the Report in terms of "accuracy and completeness."

(f) Finally, **Government Focal Points** provide the IPCC with a list of national experts, as well as comments on the scientific components of Report drafts.

All of this suggests the IPCC is transparent to potential authors and contributors about the different roles in the collaboration, and what sort of credit one is due depending on the role one is selected for. Not only this, the IPCC provides clear and detailed descriptions and responsibilities of each role, all of which are publicly available (see IPCC 2013a). The IPCC's authorship practices meet the condition of fair credit attribution by articulating what each role entails and what sort of credit is rewarded. This helps mitigate authorship disputes, such as bargaining for credit, ghosting or swelling publications.

Respectful participation requires that collaborators be treated with respect from group members to participate by being able to fulfill the obligations associated with their social role in the group. In addition, collaborators in authoritative roles are key to the success of this condition because they lead by modelling good behavior and are designated to direct group interactions, ensuring that members (especially junior and/or vulnerable members) are not undermined or shut out of the

conversation by others. In general, this condition encourages participation from all collaborators, but with the acknowledgement that members occupy various social roles in the group with differential degrees of authority.

Regarding the IPCC's practices, they support respectful participation in two ways. First, the publicly available descriptions and responsibilities of each collaborator's role makes it clear what one should expect when in the capacity of their assigned role to perform their obligations as a group member. Second, when forming Chapter teams, the IPCC states that "[a] comprehensive assessment requires author teams to include a mix of authors from different regions and from developed and developing countries to ensure that reports are not biased towards the perspective of any one country or group of countries and that questions of importance to particular regions are not overlooked" (IPCC 2013b, 1). In addition to geographical diversity, the IPCC aims to balance authorship diversity by also achieving gender balance among authors, including members with and without experience working on IPCC reports, as well as members with expertise from industry and non-profit organizations. These aims work to include the participation of a diverse body of members to ensure the most comprehensive and balanced Report. Finally, the IPCC's comprehensive review process that each Working Group undergoes also promotes respectful participation. Figure 5 shows the number of comments each Working Group receives, the number of experts involved, as well as the number of governmental bodies that participate in the review process. A First Order Draft is produced, which then gets circulated to expert reviewers for comments. The First Order Draft is then revised into a Second Order Draft, which then gets circulated to the same expert reviewers, as well as some new ones, and government bodies for feedback. Lastly, a final draft is produced taking into account the last set of comments. The Working Groups also produce a Summary for Policy Makers (SPM) outlining the main conclusions from each of their respective reports. The SPM undergoes one round of review from experts and governments. For example, Working Group 1 received 21,400 comments from 659 experts working on the physical aspects of climate science on their First Order Draft. For the Second Order Draft they received 31,422 comments from 800 experts (which included both experts who gave comments on the First Order Draft and newer experts) and 26 different governments. Figure 6 is the IPCC's visual outline of the preparation of the reports.

Number of review comments on Fifth Assessment Report

		Number of comments	Experts	Governments
Working Group I	First Order Draft	21,400	659	-
	Second Order Draft	31,422	800	26
Working Group II	First Order Draft	19,598	563	-
	Second Order Draft	28,544	452	33
Working Group III	First Order Draft	16,169	602	-
	Second Order Draft	19,554	444	24
Synthesis Report	First Order Draft	5,944	85	42
Total		142,631	-	-

Note: some experts register for more than one Working Group and the Synthesis Report

Figure 5: IPCC chart highlighting the review process (IPCC 2015).

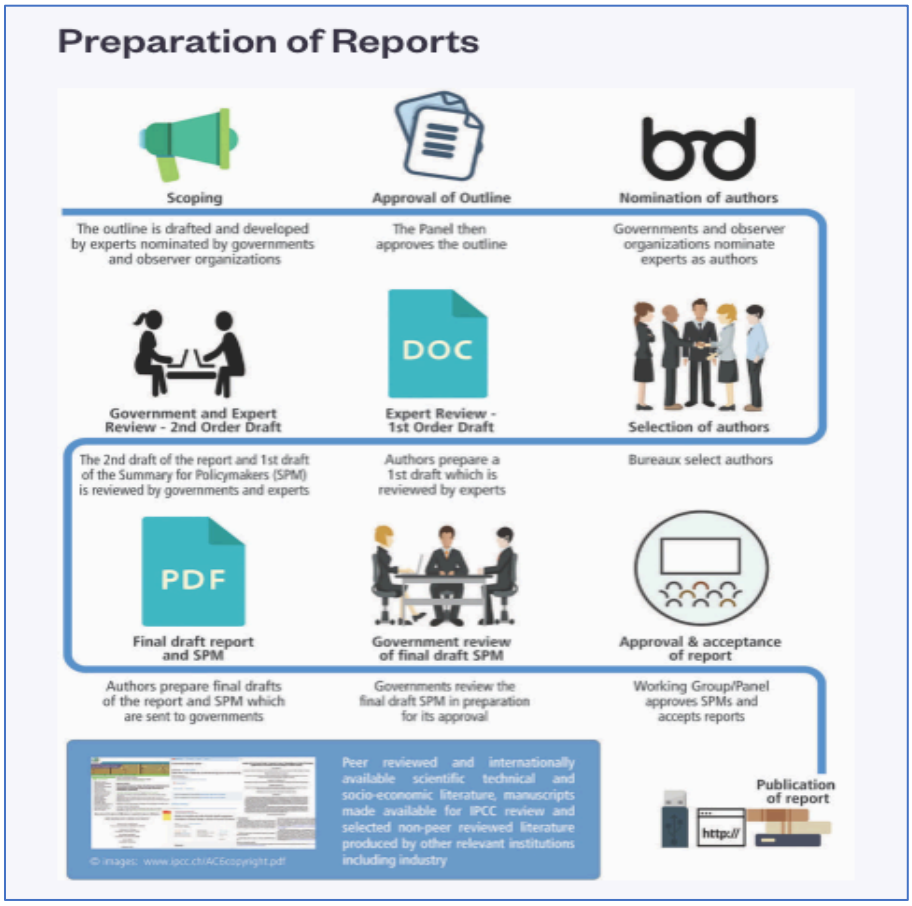

Figure 6: Process of preparation of reports (IPCC 2020a).

This extensive review process supports respectful participation by incorporating a variety of different perspectives into each of the Working Group reports and the SPM report. Experts and governments are encouraged to register as reviewers to facilitate "the participation of experts encompassing as wide a range of views, expertise and geographical representation as possible" (IPCC 2015, 1). Moreover, comments received are addressed by authors, and written responses are provided (2015, 2). Reviewers and government bodies have a central role in the IPCC reports because they help make the documents more comprehensive by putting forth a range of considerations.[112]

The other two conditions, group leader accountability and external grievance reporting, can be met by looking at the IPCC's "Code of Conduct for Meetings and Events." The code of conduct broadly covers "discrimination, harassment, which includes sexual harassment, and abuse of authority" (IPCC 2019, 1). There is a zero-tolerance approach to code of conduct violations for all IPCC members:

> IPCC meetings and events are professional, respectful and harassment-free environments for all participants. Participants are expected to conduct themselves with integrity and in a manner that is professional, respectful, tolerant and responsible. A zero-tolerance approach will be applied to any form of discrimination or harassment, including sexual harassment and bullying, at IPCC meetings and events. (IPCC 2019, 1)

Recall the **group leader accountability** condition aims to prevent group leaders from abusing their authority, especially upon members in junior or vulnerable social roles. The IPCC's code of conduct explicitly states that abuse of authority is not tolerated. In particular, they define abuse of authority in a manner that recognizes how individuals in influential roles can use their power to leverage others. The UN Secretary-General Bulletin's definition of abuse of power is adopted, and described as follows:

> Abuse of authority is the improper use of a position of influence, power or authority against another person. This is particularly serious when a person uses his or her influence, power or authority to improperly influence the career or employment conditions of another, including, but not limited to, appointment, assignment, contract renewal, performance evaluation or promotion. Abuse of authority may also include conduct that creates a hostile or offensive work environment which includes, but is not limited to, the use of intimidation, threats, blackmail or coercion. Discrimination and harassment, including sexual harassment, are particularly serious when accompanied by abuse of authority. (Ban 2008, 2)

[112] The IPCC describes the review process as being governed by three principles:

> Three principles governing the review should be borne in mind. First, the best possible scientific and technical advice should be included so that the IPCC Reports represent the latest scientific, technical and socio-economic findings and are as comprehensive as possible. Secondly, a wide circulation process, ensuring representation of independent experts (i.e. experts not involved in the preparation of that particular chapter) from developing and developed countries and countries with economies in transition should aim to involve as many experts as possible in the IPCC process. Thirdly, the review process should be objective, open and transparent. (IPCC 2013a, 6)

While the code of conduct broadly expects all participating members of the IPCC to treat each other with respect and integrity, it also highlights differential power dynamics that emerge from there being different social roles. IPCC members in authoritative roles, in particular, should not use their authority for unethical purposes.

The **external grievance reporting** condition protects junior and socially vulnerable collaborators by providing an independent avenue to report grievances such as coercion, abuse of power, or inappropriate behavior. The IPCC's code of conduct offers an avenue for members to confidentially report misconduct via email, and a point of contact person handles the case. Participants also have the option of speaking to someone in an authoritative position at the IPCC (assuming they are not the ones involved). In addition, the policy ranges over a variety of interactions, not only those limited to in-person interactions, such as e-mail and teleconferences (IPCC 2019, 1). The code of conduct states:

> In the event that a participant feels like an issue has arisen regarding disrespectful treatment, harassment or discrimination, regardless of whether it rises to the level of breach of this conduct, the participant is encouraged to speak to email about it in confidence. A designated point of contact will be identified from the TSU, but if they feel more comfortable, participants could discuss with a Co-Chair or the IPCC Deputy Secretary.
>
> All reports and allegations of breaches to this code of conduct will be handled sensitively and in confidence. The immediate priority will be to protect the privacy of person(s) exposed to the breach of conduct and the need for further action will be assessed carefully on a case-by-case basis. Note that this policy applies not only to in-person meetings, but also to teleconferences, e-mail exchanges and other interactions. (IPCC 2019, 1-2)

Grievances are handled confidentially and independently. The public documentation of an explicit code of conduct makes explicit what sort of behavior is acceptable among IPCC members.

I want to now mention two places in which the IPCC's procedures and policies can be improved. In particular, the ones that cover group leader accountability and external grievance reporting. For the former, even though the IPCC recognizes that individuals in positions of authority can abuse power afforded to their roles, little is said about how this behavior will be mitigated. It would be an improvement if they also insert a body that occasionally assesses the leadership skills of leaders to monitor and prevent any actions related to abuse of power. So not only do IPCC members know that abuse of power is not tolerated, they also know that group leaders are held to a high standard coupled with occasional assessments. For external grievance reporting, depending on the nature of the report, grievances should be handled by an independent body of peers outside the IPCC that follow outlined steps for how confidential reports should be handled. One might be worried that if misconduct grievances are handled internally by individuals with professional and interpersonal connections to the parties involved, then their assessments may not be sufficiently independent. This is also important because groups who are against policies in favor of climate change are constantly finding ways to undermine the IPCC's respectability as an authoritative expert body. Thus assigning an independent body to handle

misconduct would help promote a fair grievance procedure for the parties involved, and it would mitigate complaints that the IPCC manages misconduct internally to hide transparency.

Overall though, the assessment of the IPCC's procedures and policies using the four normative criteria for scientific collaboration I proposed reveals that the IPCC is a well-coordinated and thoughtful body, sensitive to a variety of social dynamics in group settings. They generally satisfy the conditions for well-functioning collaboration because they have procedures and policies in place to address social dynamics that can make their collaboration vulnerable, though there is room for improvement. The policies are generally well-defined and transparent, and also publicly accessible. The IPCC's explicit statements about practices and code of conduct make clear what is expected of IPCC members, and what counts as a violation of group norms. In the next section, I briefly return to answering climate change skeptics.

IV Answering the Skeptics

We now have resources to offer a positive answer to climate change skeptics for why the IPCC's consensus position on anthropogenic climate change is trustworthy. First, let me start with some quotes from skeptics to remind us of the flavor of their criticisms. Here is one quote from the *Heritage Foundation* claiming that climate scientists suppressed dissent, revealed by the 1000 leaked emails from the "Climategate" scandal, which mentions IPCC work:

> Hackers leaked thousands of e-mails and other documents from the University of East Anglia's Climate Research Unit that detailed how these climatologists, many with important roles in promulgating the official U.N. science, refused to share data, plotted to keep dissenting scientists from being published in leading journals, and discarded original data. (Loris 2010)

And here is a more recent one that claims there is groupthink about climate change:

> [D]issenters were ignored. And for nearly 20 years the "consensus" rolled on, ever more extreme in its apocalyptic claims, with each new IPCC report scarier than the last. By 2006, Al Gore's *An Inconvenient Truth* was outdoing them all.
>
> Anyone daring to question the "consensus" was now being vilified as just an "anti-science denier," no better than those crazies who deny the reality of the Nazi Holocaust.
>
> […] Trump was thus finally calling the bluff of the groupthink which for 30 years had driven the whole global warming scare. If other Western countries wanted to commit economic suicide, that was their affair. But the rest of the world was no longer taken in by it, and the US was now with them. (Booker 2018)

I previously argued that aggregative consensus and the threshold that trustworthy and legitimate consensus positions be nearly or completely unanimous should not be employed to assess the IPCC's consensus position (see Chapter 3). The IPCC group is engaged in collaboration involving a sophisticated

division of cognitive labor among three Working Groups with research experts from a variety of subfields, as well as governmental bodies, non-profit organizations, and industry experts. No single individual in the group can assess all of the empirical evidence considered in the report. Given this, aggregative consensus and unanimity are not relevant considerations. Furthermore, as I argued in Chapter 4, what makes collaborative efforts distinct has to do with the relations of dependence members share with one another to realize the group's aims. In particular, members occupy a variety of social roles in the group, and their expectations of each other are based on the obligations associated with that role. Authority structures are key here since different social roles confer different degrees of authority and so members are not in equal standing to each other. This provides a further reason to reject using aggregative consensus to assess collaboratively produced consensus because members are organized hierarchically, not democratically. Collaboratively produced consensus should thus be assessed using a different set of standards. We should focus on the social and organizational structure of the collaboration itself. We should look at, not just whether collaborations have robust evidence for their claims, but importantly whether they are organized in a way that mitigates negative social dynamics, along the dimensions of the four normative conditions (credit, respectful participation, group leader accountability, and external grievance reporting) I proposed. And if negative social dynamics do arise, the group ought to have structural resources to adequately handle them. My analysis from the last section showed that the IPCC's procedures and practices for collaboration more or less cover the normative conditions I outlined.

In sum, we should first make it clear that the standards for aggregative and collaborative consensus are distinct. The skeptics' appeal to aggregative consensus is misguided when assessing the trustworthiness of the IPCC's consensus position. When assessing collaborative consensus, we should look at the group's procedures and practices that manage members' relations with one another for the group to function well. The IPCC's consensus position should be trusted because, not only do they have robust empirical evidence on their side, but also their procedures and practices meet conditions for well-functioning collaboration, which help mitigate social dynamics like groupthink.

Conclusion

In this final chapter, I offered four working criteria for well-functioning collaboration: fair credit attribution, respectful participation, group leader accountability, and external grievance reporting. Together these criteria help mitigate some vulnerabilities of socially cohesive collaboration, such as groupthink and abuse of power. I argued that when looking at whether to trust the consensus position of a collaborative group, these proposed criteria focus on the relevant features regarding whether the group is a well-functioning one. I then used these criteria to assess the IPCC collaboration. I argued that we have good reason to trust their consensus position on anthropogenic climate change because their collective procedures and practices generally meet the norms outlined, though some areas could be strengthened. Given this, we can tell skeptics that they have incorrectly invoked aggregative consensus

and the unanimity criterion to assess the IPCC's consensus position. The IPCC's consensus position is a collaboratively produced one, and their collaboration satisfies norms for well-functioning collaboration.

Bibliography

Abritis, Alison, Adam Marcus, and Ivan Oransky. n.d. “Retraction Watch.” Retraction Watch. https://retractionwatch.com.

Albert, Tim, and Elizabeth Wager. 2003. “How to Handle Authorship Disputes: A Guide for New Researchers.” Committee on Publication Ethics. https://doi.org/10.24318/cope.2018.1.1.

Andersen, Hanne. 2014. “Co-Author Responsibility.” *EMBO Reports* 15 (9): 914–18.

Ban, Ki-Moon. 2008. “Prohibition of Discrimination, Harassment, Including Sexual Harassment, and Abuse of Authority.” ST/SGB/2008/5. Secretary-General’s Bulletin. United Nations Secretariat. https://undocs.org/en/ST/SGB/2008/5.

Beatty, John. 2017. “Consensus: Sometimes It Doesn’t Add Up.” In *Landscapes of Collectivity in the Life Sciences*, edited by Snait Gissus, Ehud Lamm, and Ayelet Shavit, 179–89. Cambridge, MA: MIT Press.

Beatty, John, and Alfred Moore. 2010. “Should We Aim for Consensus?” *Episteme* 7 (3): 198–214.

Biagioli, Mario. 2003. “Rights or Rewards? Changing Contexts and Definitions of Scientific Authorship.” In *Scientific Authorship: Credit and Intellectual Property in Science*, edited by Mario Biagioli and Peter Galison, 253–80. New York: Routledge.

Booker, Christopher. 2018. “Groupthink on Climate Change Ignores Inconvenient Facts.” *Foundation for Economic Education*. https://fee.org/articles/groupthink-on-climate-change-ignores-inconvenient-facts/.

Brown, DeNeen L. 2017. “You’ve Got Bad Blood: The Horror of the Tuskegee Syphilis Experiment.” *The Washington Post*, May 6, 2017. https://www.washingtonpost.com/news/retropolis/wp/2017/05/16/youve-got-bad-blood-the-horror-of-the-tuskegee-syphilis-experiment/.

CASRAI. 2014. “CRediT - Contributor Roles Taxonomy.” CASRAI. 2014. https://casrai.org.

Collins, Harry, and Robert Evans. 2002. “The Third Wave of Science Studies: Studies of Expertise and Experience.” *Social Studies of Science* 32 (2): 235–96.

Committee on Publication Ethics (COPE). 2020a. “Authorship and Contributorship.” Publication Ethics. 2020. https://publicationethics.org/authorship.

———. 2020b. “Promoting Integrity in Scholarly Research and Its Publication.” Publication Ethics. 2020. https://publicationethics.org.

Cook, John, Dana Nuccitelli, and Sarah A. Green. 2013. “Quantifying the Consensus on Anthropogenic Global Warming in the Scientific Literature.” *Environmental Research Letters* 8 (2): 024024.

Goodwin, Jean. 2009. “The Authority of the IPCC First Assessment Report and the Manufacture of Consensus.” *Working Draft*.

ICMJE. 2020. “Defining the Role of Author and Contributions.” International Committee of Medical Journal Editors. 2020. http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html.

IPCC. 2013a. “Procedures for the Preparation, Review, Acceptance, Adoption, Approval and Publication of IPCC Reports.” https://www.ipcc.ch/site/assets/uploads/2018/09/ipcc-principles-appendix-a-final.pdf.

———. 2013b. “IPCC Factsheet: How Does the IPCC Select Its Authors?” https://www.ipcc.ch/site/assets/uploads/2018/02/FS_select_authors.pdf.

———. 2013c. “Principles Governing IPCC Work.” Batumi. https://archive.ipcc.ch/pdf/ipcc-principles/ipcc-principles.pdf.

———. 2015. “IPCC Factsheet: How Does the IPCC Review Process Work?” January 15, 2015. https://www.ipcc.ch/site/assets/uploads/2018/02/FS_review_process.pdf.

———. 2018. “Working Group I Contribution to the IPCC Sixth Assessment Report (AR6-WG1).” IPCC. 2018. https://archive.ipcc.ch/report/authors/report.authors.php?q=35&p=&p.

———. 2019. “Code of Conduct for Meetings and Events of the Intergovernmental Panel on Climate Change.” https://www.ipcc.ch/site/assets/uploads/2019/07/Code_of_Conduct.pdf.

———. 2020a. “Preparing Reports.” IPCC. 2020. https://www.ipcc.ch/about/preparingreports/.

———. 2020b. “Structure of the IPCC.” IPCC. 2020. https://www.ipcc.ch/about/structure/.

———. 2020c. “Working Groups.” IPCC. 2020. https://www.ipcc.ch/working-groups/.

Janis, Irving. 1972. *Victims of Groupthink*. Boston: Houghton Mifflin.

Longino, Helen E. 2002. *The Fate of Knowledge*. New Jersey: Princeton University Press.
Loris, Nicolas D. 2010. “How the ‘Scientific Consensus’ on Global Warming Affects American Business—and Consumers.” *The Heritage Foundation Backgrounder*, 2010, No. 2479 edition.
Melo-Martín, Inmaculada de, and Kristen Intemann. 2018. *The Fight Against Doubt*. Oxford University Press.
National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. 1979. “Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research.”
Oreskes, Naomi. 2004. “The Scientific Consensus on Climate Change.” *Science* 306: 1686.
———. 2019. *Why Trust Science?* Oxford: Princeton University Press.
Panter, Michaela. 2020. “Ghost Authorship.” American Journal Experts (AJE). 2020. https://www.aje.com/arc/ghost-authorship/.
Pierson, Charon. 2019. “View from COPE - Qualitative Analysis of 134 Authorship Cases.” Panel Discussion. https://publicationethics.org/files/responsible_authorship_charon_pierson_wcri_2019.pdf.
Rehg, William, and Kent Staley. 2017. “‘Agreement’ in the IPCC Confidence Measure.” *Studies in History and Philosophy of Modern Physics* 57: 126–34.
Solomon, Miriam. 2006. “Groupthink versus the Wisdom of Crowds: The Social Epistemology of Deliberation and Dissent.” *The Southern Journal of Philosophy*, no. 44 Supplement: 28–42.
Tollefsen, Deborah. 2006. “Group Deliberation, Social Cohesion, and Scientific Teamwork: Is There Room for Dissent?” *Episteme* 3 (1–2): 37–51.
Verfaellie, Mieke, and Jenna McGwin. 2011. “The Case of Diederik Stapel.” News. *American Psychological Association*. https://www.apa.org/science/about/psa/2011/12/diederik-stapel.
Wylie, Alison. 2006. “Socially Naturalized Norms of Epistemic Rationality: Aggregation and Deliberation.” *The Southern Journal of Philosophy*, no. 44 Supplement: 43–48.

Concluding Remarks

This dissertation aims to offer an analysis of scientific consensus with attention to contemporary collaborative practice. The dynamics of collaborative practice turn out to be centrally informative if we want to understand consensus that, to use John Beatty's words again, "sometimes doesn't add up" (2017).

In Chapter 1, I distinguished between two broad categories of aggregative and collaborative consensus classifying some models in the philosophical literature according to these categories. I noted that conceptually aggregative and collaborative consensus models map onto different types of consensus in scientific practice oriented for different purposes. The standards of assessment for aggregative and collaborative consensus are also distinct. One central aim of the dissertation is to offer norms to assess socially cohesive groups that produce collaborative consensus. In Chapter 2, I engaged with Fagan's recent critique of Gilbert's joint commitment model (a collaborative consensus model) for capturing scientific consensus. I argued that a proponent of Gilbert's model can adequately address Fagan's concerns, especially when considering it in the context of collaboration. However, this does not mean that Gilbert's model is the best one. I argued that we ought to keep the various collaborative consensus models available as resources, since collaborations can vary in terms of social dynamics and organizational structure. No single collaborative consensus model has all of the conceptual resources to articulate collaborative groups in practice.

In Chapter 3, I applied my distinction and analyses to a real-world case: the climate change controversy, looking at arguments from skeptics, and a prominent defender of the IPCC's consensus position, Naomi Oreskes. I argued that they both invoke aggregative consensus when discussing the IPCC, which is a mistake. My analysis showed that the IPCC is structured as a collaboration, and their group position amounts to collaborative consensus. We should instead assess the IPCC as a collaboration. In Chapter 4, I argued that in contemporary scientific collaborations, group members are highly dependent on one another in terms of expertise, as well as the expectations and obligations associated with assigned social roles. Moreover, these groups are often hierarchically organized since social roles confer differential degrees of authority. When considering norms to assess collaborative groups, dependence relations should be a central consideration.

Finally, in Chapter 5, I began the discussion by pointing out that the social cohesion of collaborative groups can make them susceptible to negative social dynamics, such as abuse of authority, or groupthink dynamics that suppress dissent. I proposed four normative criteria for well-functioning collaboration to mitigate these sorts of dynamics (fair credit attribution, respectful participation, group leader accountability, and external grievance reporting). When looking at publicly declared consensus positions from collaborations, like the IPCC's, one way we can deem them trustworthy, or not, is based

on whether they incorporated social procedures and practices that deter toxic dynamics where members agree to let a position stand as the group's view out of pressure from others. Given this, we should evaluate whether the IPCC has social procedures and practices that could mitigate the negative social dynamics of collaboration discussed, in addition to satisfying other norms of scientific inquiry like responsiveness to empirical data. My assessment showed that, though there is some room for improvement, the collaboration generally satisfies the proposed normative criteria for well-functioning collaboration.

The issues considered in this dissertation have also raised a number of additional questions, some of which I hope to explore in future work. One is the development of the proposed normative criteria for collaboration. The fair credit attribution norm, for example, would benefit from further analysis of other co-authorship practices, such as publication retractions. When a co-authored publication gets retracted, which individuals are held responsible and how are junior collaborators in particular affected? Are there policies in place that protect junior collaborators, and what happens to their research trajectory long-term? Differently, the proposed normative criteria would benefit from the feedback of scientific practitioners at various career stages, including those who assess grant applications. These criteria would only be useful if there is sufficient receptiveness and uptake from actual practitioners and the institutions that govern their work. Another avenue to consider is what it means to have inclusive collaborations that produce consensus positions, along a variety of different social and epistemic dimensions. For example, would the consensus positions of collaborations that include the expertise of a variety of different communities that go beyond scientific experts be considered more trustworthy? This would require more engagement with STS literature and could potentially offer new ways to strengthen relations between experts and lay public communities.

Overall, my hope is that the contents of this dissertation could offer strategies for communicating to various public communities about the complexities of consensus statements from expert bodies like the IPCC. An implicit consideration for me has been that many members of public communities are not in a position to assess the evidential content behind the IPCC's consensus position. In addition to aggregative consensus being the wrong type for evaluating the IPCC's position, emphasis on it by defenders of the IPCC may simply look like an appeal to authority that, as Goodwin puts it, "puts the audience in a position such that they will appear imprudent if they conspicuously go against the view of someone who knows more" (2009, 3). But simply saying that there is consensus about anthropogenic climate change (or vaccines, or evolutionary theory, or COVID-19) among the relevant experts is not enough to convince people. Matters are even worse when we acknowledge that scientific communities have a fractured relationship with many public communities, especially marginalized communities (e.g., racist, ableist, and sexist science). A lot of cachet has been given to science's prestige as *a*, if not *the*, most successful knowledge producing discipline. Injustices committed in the name of productive science are often relegated to the margins of "bad" science. But these injustices have themselves sewn discord, distrust, and ambivalence about science. A top-down approach will therefore not work in getting people

to trust the experts – it will require more than asserting authority of knowledge. Here I'm following Goodwin's suggestion, who writes: "Whereas non-experts almost by definition are unable to assess an expert's reasoning […], they may be well capable of judging social facts, such as whether some procedures were inclusive" (Goodwin 2009, 3). The IPCC and other expert bodies would, I think, be in a better position if they highlight the strengths of their social practices and procedures that govern their collaborative effort, since these are social features that non-experts could understand.

Bibliography

Beatty, John. 2017. “Consensus: Sometimes It Doesn’t Add Up.” In *Landscapes of Collectivity in the Life Sciences*, edited by Snait Gissus, Ehud Lamm, and Ayelet Shavit, 179–89. Cambridge, MA: MIT Press.

Goodwin, Jean. 2009. “The Authority of the IPCC First Assessment Report and the Manufacture of Consensus.” *Working Draft*.