Judgement aggregation in scientific collaborations: The case for waiving expertise*

This is the pre-corrected version of

https://doi.org/10.1016/j.shpsa.2020.08.001

In Press, *Studies in History and Philosophy of Science*

Alexandru Marcoci[†] James Nguyen[‡]

Abstract

The fragmentation of academic disciplines forces individuals to specialise. In doing so, they become experts over their narrow area of research. However, ambitious scientific projects, such as the search for gravitational waves, require them to come together and collaborate across disciplinary borders. How should scientists with expertise in different disciplines treat each others' expert claims? An intuitive answer is that the collaboration should defer to the opinions of experts. In

*Authors are listed in alphabetical order. We would like to thank Luc Bovens, Harry Collins, Christian List, Graham Oddie, Meghan Page and multiple anonymous referees for their comments on earlier versions of this paper. We are grateful to audiences at the 2017 annual British Society for the Philosophy of Science conference in Edinburgh, and the 2018 American Philosophical Association Central Division meeting in Chicago. JN acknowledges the Jacobson Committee at the University of London for postdoctoral support, and the British Academy , for a Rising Star award on Epistemological Pluralism.

[†]Department of Philosophy, University of North Carolina at Chapel Hill and Centre for Philosophy of Natural and Social Science, LSE, marcoci@unc.edu.

[‡]Institute of Philosophy, University of London, Department of Philosophy, UCL, and Centre for Philosophy of Natural and Social Science, LSE james.nguyen@sas.ac.uk.

this paper we show that under certain seemingly innocuous assumptions, this intuitive answer gives rise to an impossibility result when it comes to aggregating the beliefs of experts to deliver the beliefs of a collaboration as a whole. We then argue that when experts' beliefs come into conflict, they should waive their expert status.

Keywords: scientific rationality, judgement aggregation, social epistemology, collaboration in science

1 Introduction

Academic disciplines are increasingly fragmented, and this naturally leads to diverse areas of expertise within them. However, ambitious research projects, such as LIGO-Virgo[1] and ATLAS[2] require experts from different fragmented disciplines come together and make collective decisions on scientific matters. But if the doxastic state of a collaboration as a whole is supposed to depend in a certain way on the expert beliefs of its members, then we arrive at an impossibility result similar to Sen's Liberal Paradox (1970). Just as Sen's Lewd and Prude were forced into inconsistency by their decisiveness over their own personal spheres, a scientific collaboration may be forced into inconsistency if experts are taken to be decisive over their respective areas of expertise. At least this will be the case if we require that what the collaboration as a whole tells us about the world be generated by a suitably-constrained function on the beliefs of its constitutive members. We show this by importing Dietrich and List's (2008) result from the context of aggregating individual judgements to that of aggregating scientific expert judgements. If this is the correct way of thinking about what scientific collaborations tell us about the world, it seems they cannot protect themselves from inconsistencies. Building on the ideas of Gibbard (1974) we suggest that the best way of avoiding inconsistencies is for experts to contribute beliefs on a par with everyone else's. As such we argue against the hegemony of experts in scientific collaborations.

The paper is structured as follows. We begin by outlining the notions of collaboration, fragmentation, and expertise we have in mind and discuss how scientists

[1] `https://www.ligo.org/`.

[2] `https://atlas.cern/discover/collaboration`.

from different fields may try to come to a collective decision on what the collaboration they are a part of tells us about the world (Section 2). We then present the formal model of judgement aggregation, and, Dietrich and List's impossibility result, along with an interpretation in terms of aggregating the beliefs of scientists (Section 3). We incorporate the idea of waiving expertise into the model and prove a possibility result (Section 4), and argue for its plausibility in the contexts under consideration (Section 5).

2 Collaboration and fragmentation in science

In this section we discuss the notions of collaboration in science, and the fragmentation of scientific disciplines. We then discuss how these come together in cases of interdisciplinary collaboration.

2.1 Collaboration

Scientific collaborations have become the new norm in science. This is already a common place remark in the medical sciences. For instance, Cronin (2001, p. 561) noted that the average number of authors per article in the *Journal of Neurosurgery* and *Neurosurgery* increased from 1.8 in 1945 to 4.6 in 1995 and Constantian (1999) found that over the last century, the percentage of single-authored papers published in the *New England Journal of Medicine* (and its precursor) shifted from 98% to less than 5%, with the mean being 6 co-authors. The same pattern has been noted in other disciplines. A survey of 9.9 million research papers from the Institute for Scientific Information (ISI)'s Web of Science database and of 2.1 million patent records, covering the fields of science and engineering since 1955, social sciences since 1956, arts and humanities since 1975 and all

U.S. registered patents since 1975 (Wuchty et al., 2007) has found a systematic and general rise in the number of collaborations. In science and engineering, average team sizes have gone from 1.9 in 1975 to 3.5 in 2000. In the social science, 17.5% of the papers were collaborative in 1955, whereas by 2000 the share of collaborative research outputs rose to 51.5% (most of them due to pairs of researchers). For patents, average team size has gone from 1.7 in 1975 to 2.3 inventors per patent in 2000. The same trend was observed in the arts and humanities, however, it is still the case that the vast majority of research produced here is the work of a single individual (90%).

A second trend is that collaborations have become more geographically dispersed. Jones et al. (2008) reviewed 4.2 million research papers in the Web of Science database, produced at 662 U.S. universities between 1975 and 2005, covering science and engineering, social sciences and arts and humanities. They found that the share of between-school collaborations in science and engineering quadrupled between 1975 and 2005 to 32.8%. The share of social science papers written by geographically dispersed collaborations also increased over the same time period and reached 34.4%. Overall, multi-site collaborations increased in 98% of the subfields of engineering and science studied, in all the subfields in social science and in 67% of the subfields in arts and humanities. The same trend towards more dispersed collaborations can be seen when looking at international collaborations. Coccia and Wang (2016) found that 50% of co-authored papers in clinical medicine in 2012 involved international collaborations, whereas the share was 64% in physics and 86% in astronomy.

Alongside the shift from single-authored to co-authored and from single-site to multi-site collaborations, a third significant trend in the way science is conducted is from small groups to large scientific collaborations. Big science (Price, 1963), as this phenomenon is known, "is distributed widely over time and space, involves researchers from multiple disciplines, and typically results in publications with dozens of listed authors" (Huebner et al., 2018, p. 95). Indeed, in 1981 there was only 1 paper with more than 100 authors; by 1994 this number increased to 182 (McDonald, 1995). And some papers, e.g. (Abbott et al., 2016), have over 1000 authors. Writing about the large detector teams surrounding colliding beams accelerators, Galison (2003, p. 325) quips that "it hardly takes algebraic topology to reckon that quite soon only a handful of these will embrace the careers of nearly all of the seven thousand experimental particle physicists expected to be employed at the end of the century".

For the most part, the emergence of collaborative research is seen as advantageous for the scientists. For instance, Wray (2002) argues that collaborations: tend to produce more successful papers, measured by looking at the number of citations (cf. Wuchty et al., 2007; Jones et al., 2008); allow for the pursuit of more ambitious research questions that couldn't be answered from within a single field of expertise, no matter how wide that may be (cf. Huebner et al., 2018); make it easier to access past relevant achievements - even if a team member is unaware of a relevant result, someone else in the team might remember it; foster productivity; and finally, that they provide training opportunities for junior researchers.

2.2 Fragmentation

Most, if not all, of modern academic disciplines exhibit a large degree of fragmentation and specialisation. A first level of fragmentation is provided by ISI that divides science into three main categories: science and engineering, social science, and humanities. An astronomer working on the identification of gravitational waves may have very little to say to a political scientist calculating the voting power of different groups in the EU Parliament. The same observation applies, however, if we 'zoom in': the expertise of a practising physicist does not overlap with the expertise of a doctor conducting randomized control trials to test the efficacy of a newly developed medicine. This also provides our second level of specialisation: ISI's classification system further divides science and engineering into 171 subfields, the social sciences into 54, and the arts and humanities into 27 (Wuchty et al., 2007). And we needn't stop there. The same considerations apply when we 'zoom in' further, to particular areas of research of the sort an international conference might be dedicated to: a chemist synthesising metal nanoparticles is far removed from one synthesising artificial molecules capable of transmitting genetic information. At each level of grain we find fragmentation accompanied by expertise. Obviously, this will not hold all the way down. Our point is that it holds for many areas of research which are still considered to be the same discipline.

For the most part, the fragmentation of any field into sub-disciplines is seen as a desirable phenomenon to be expected. In a recent editorial, Casadevall and Fang (2014) argue that fragmentation in science occurs naturally as a consequence of a need for increased efficiency. They give the example of the field of 'Infectious Diseases'. In 1978 the president of the Infectious Diseases Society of America said that he "cannot conceive the need for 309 more infectious disease experts unless they spend their time culturing each other" (p. 1356). However, as Casadevall and Fang note, as a consequence of the AIDS epidemic, the number experts in infectious diseases started growing and "there are now 7,500 board-certified infectious disease specialists" (ibid). Moreover, a new even more specialised subfield emerged within the infectious diseases community focusing on HIV alone and organising itself in the HIV Medicine Association.

2.3 Collaboration over fragmented fields

Now, if we want to know what a large scientific collaboration spanning multiple fragmented fields of study tells us about the natural world, then we need a way of combining the beliefs of experts coherently. For any suitably cross-disciplinary collaboration, if we want to know what the collaboration tells us about the world, we need to combine the beliefs of experts working in different fields as part of the joint research project. How should we combine these beliefs to deliver what a collaboration of scientists as a whole tells us? Here are some, seemingly innocuous, requirements that such a combination procedure should meet.

First, the doxastic state of a collaboration should supervene on the state of its constitutive members: the state of the collaboration is a function on the doxastic states of the individuals involved. Consider, as a toy example, a case involving two scientists interested in the effect of carbon dioxide emissions on ocean temperature (adapted from Dietrich and List, 2008, p. 60). Suppose that one scientist is an expert on carbon dioxide emissions (and consider, e.g. the proposition p: carbon dioxide emissions are above a critical threshold). Assume they are one of the scientists working on measuring the level of CO_2 in the atmosphere using data generated by spectrometers. Suppose the other is an expert on ocean temperature (consider, e.g. the proposition q: the oceans are warming), and suppose that they need to supply a document reporting their joint belief on the topic at hand to policy makers indicating their findings. The idea we're working with here is that what they report, or what they believe, with respect to p, q, and other more complex propositions involving them (including how they relate to each other and other relevant propositions) is a function on the doxastic states of the individuals with respect to the relevant propositions in question.

It is worth highlighting that this is a relatively minor requirement, and is consistent with a variety of different metaphysical interpretations. It neither precludes, nor demands, that the doxastic state of the collaboration is ontologically distinct from the doxastic states of the individuals involved. It also neither precludes, nor demands, that the joint doxastic state be 'reducible', in any loaded way, to the doxastic states of the individuals involved.[3]

Note also that, whilst we talk about the 'belief' of the collaboration, this is shorthand for whatever is the right way of characterising the doxastic state of the collaboration. For example, Wray (2001, 2018), in response to Gilbert (2000), argues that the appropriate doxastic state of a collaboration is 'group acceptance' rather than 'belief', since when a group accepts a proposition they do so voluntarily, whilst beliefs are acquired involuntarily, and a believer is compelled to believe what they think is true. This is compatible with the idea that the doxastic state of the collaboration, however it is to be characterised, nevertheless supervenes on the beliefs of the individuals (in fact,

[3]For example, it is compatible with the idea that talk of 'group belief' is paraphrase for talk of the doxastic states of the individuals; compatible with the idea that a group can believe a proposition not believed by anyone in the group; and also compatible with the idea that scientific collaborations are doxastic group agents in just the same way that individuals are doxastic agents). Thus, we take this to be relatively uncontentious. But of course there are exceptions see Bird (2010, 2014).

it's also compatible with the idea that it supervenes on whatever the relevant doxastic states of the individuals are, regardless of whether they are beliefs or characterised in terms of acceptance).[4]

Second, experts should determine what the collaboration believes about their area of expertise. In other words experts should act as local 'dictators' over the topics they have expertise in. In the toy example above, if we suppose that the first scientist is an expert on carbon dioxide emissions, e.g. they have led a decades-long research programme measuring the level of carbon dioxide in the atmosphere, they are tasked by international agencies with providing data concerning emissions by different countries and so on, then we might give them the authority to determine what the group reports concerning carbon dioxide emissions. And if we suppose that the second scientist has lead another decades-long research programme tasked with measuring ocean temperatures, then we might give them the authority to determine what the group reports concerning ocean temperatures.

In the literature on opinion aggregation in democratic societies, this view is typically

[4]One may question whether what a collaboration tells us about the natural world relies only on what individual scientists believe at a time. In the philosophy of science literature, it has been proposed that the way in which theories about the natural world are developed and adopted by the relevant scientific community is influenced by institutions, history, and power relations (for some intriguing examples, see Okruhlik, 1994). We grant these influences on what science tells us about the world. However they do not feed into the project we are interested in here. Rather, they arise further upstream and concern how individual scientists form their beliefs. Once this is fixed (regardless of how), the question we are interested in arises.

viewed negatively for obvious reasons.[5] However, it is not obvious the same should hold for societies in which individuals may have a genuine claim at being more likely to be closer to the truth than the rest. Indeed, "a dictatorship is usually ruled out in the social perspective, but it may be acceptable from an epistemic perspective, if that individual is by far the most competent member" (Martini and Sprenger, 2018, p. 180).[6]

And finally, whenever everyone agrees on something, the collaboration as a whole should agree on that too. We take this to be a relatively innocuous assumption about the doxastic behaviour of any group aiming to arrive at a collective decision (although we will comment on it further below).

However, results from Dietrich and List (2008) building on the liberal paradox introduced by Sen (1970), show that if these three principles hold, then we cannot guarantee that the scientific collaboration as a whole will be free of inconsistency. As an informal way of seeing how such a result can be generated consider the following situation. The first scientist in our toy example believes that carbon dioxide emissions are over the critical threshold (p); if they are over that threshold then ocean temperatures are warming ($p \to q$); and that ocean temperatures are warming (q). The

[5]Note that the notion of expertise is taken to be restricted to only certain areas of the collaboration, rather than the entirety of what the group reports, the latter of which corresponds to the notion of a typical 'dictator' in the judgement aggregation literature.

[6]To further motivate the *prima facie* plausibility of this principle in the case of scientific collaborations consider the following remark by Nobel prize winning physicist Carlo Rubbia: "I cannot neglect the fact that people who are working on [those aspects of the experiment under discussion] have more weight than people who aren't" (Bright et al., 2018, quoted at pp. 242-3).

second scientist believes that it's not the case that carbon dioxide emissions are over the threshold ($\neg q$); but if they were then ocean temperatures would be warming ($p \to q$); and that ocean temperatures are not warming ($\neg q$). As a result of our assumption that the doxastic state of the group be a function on the beliefs of the individuals in the group, the joint belief needs to be determined in some way from the beliefs of the individuals. As a result of our assumption about expertise, the joint belief contains the propositions p and $\neg q$ (since the first scientist is an expert on carbon dioxide emissions and believes that they are higher than the critical threshold, and the second an expert on ocean temperatures and believes that they are not rising). And as a result of our assumption about propositions on which everyone agrees, the conditional $p \to q$ will also be part of the joint report. Thus, the joint belief will include the propositions p, $p \to q$, and $\neg q$, an inconsistency. As we will see below, the result generalises: in any collaboration across suitably fragmented disciplines, there will be cases where individually consistent experts will, collectively, deliver inconsistent claims. And we take it for granted that a collaboration supporting an inconsistent collective belief would be required to correct it. The question is, can this be done whilst still respecting the scientists' epistemic advantage over their respective areas of expertise?

3 The formal framework

In this paper we adopt a formal model for aggregating judgements over sentences, supplied by Dietrich and List (2008).[7] In this section we outline the model and provide

[7]See also List and Pettit (2002); Dietrich (2006); Dietrich and List (2007a, 2013). For a useful introduction to the framework see Grossi and Pigozzi (2014). See also Bright et al.

an interpretation in terms of aggregating the beliefs of scientists working in a collaboration spanning multiple (fragmented) disciplines. Begin by considering a group of *individuals*, $N = \{1, \dots, n\}$ with $n \geq 2$, and a language $\mathfrak{L}$. A set of sentences from $\mathfrak{L}$ is *consistent* if and only if it has a model in the specified logic.[8] For instance, $\{p, p \to q, \neg q\}$ is inconsistent in classical propositional logic, and $\{p, q\}$ is not. With respect to the intended use of this (idealised) model, we will take the set of individuals to supply the set of scientists working in a collaboration spanning multiple fragmented disciplines and $\mathfrak{L}$ to consist of all the meaningful claims common to the collaboration. Individuals in our model believe sentences in an *agenda*. An *agenda* is a finite nonempty set $X \subseteq \mathfrak{L}$ closed under negation.[9] Define a *position* on $\varphi \in X$ as either φ or $\neg\varphi$. Then

(2018) for a relevant discussion of how it can be used to model scientific collaboration. An alternative formal framework which might be employed to make sense of the cases we are considering in this paper, but which, for reasons of space, we do not consider here is the logic of belief-merging under integrity constraints see e.g. Konieczny and Pérez (2011). There are clear relationships between the two formal frameworks, see Pigozzi (2016) for discussion, and it is worth noting that the aggregation procedure we discuss in Section 4, whilst formalised in the judgement aggregation framework, shares the 'egalitarian' motivation that is present in the belief-merging approach.

[8]We restrict our focus to monotonic logics, i.e. ones in which any subset of a consistent set of sentences in the logic is consistent.

[9]More accurately, we understand an agenda as consisting not of sentences, but rather equivalence classes of sentences modulo logical equivalence. As a consequence φ and $\neg\neg\varphi$ are treated as the same element of X. For simplicity, we restrict our focus to finite agendas throughout but see fn.17.

call an agenda X *connected* if and only if for any two sentences φ and ψ, there is a set of sentences $Y \subseteq X$ such that some position on φ and some position on ψ are each individually consistent with Y, but jointly inconsistent with Y. An agenda is a set of claims a group of scientists are attempting to reach a collective decision on, and we assume that it is connected.

Although scientific fields are fragmented, the fragments are not logically disconnected from one another. We think that this holds even at the most general level. Interdisciplinary work shows that the classical categories overlap, and the literature on inter-theoretic reduction supports the thesis that an overarching agenda consisting of sentences from each of these categories would remain connected (Dizadji-Bahmani et al., 2010). Those uncomfortable with the general claim should note that the connectedness of agendas becomes even more plausible when we consider collaborations that work together towards a common scientific goal such as the detection of a gravitational wave. There, questions of optics, seismology, astronomy, and big data are relying on one another to deliver a coherent picture of nature.

The *belief set* of an individual i is a consistent set $A_i \subseteq X$. A profile of belief sets is an n-tuple of belief sets $(A_1, \ldots, A_n)$. When it comes to forming a collective judgement, we need a *judgement aggregation function* F that takes profiles to belief sets. The input of the function are the sequences of belief sets, with each belief set being identified with the beliefs of an individual scientist with respect to sentences in the agenda. The value of F is the set of sentences agreed upon by the collaboration as a whole: the joint belief, or report, for that sequence of individual beliefs.

Note that since we are defining a belief set as a consistent (but not necessarily complete, in the sense defined below) set of sentences from the agenda, we're allowing both individuals and groups to refrain from taking positions on some sentences in the agenda. However, by requiring the codomain of an aggregation function be the set of consistent subsets of X, we are only considering aggregation procedures that are 'resolute' in the sense that each profile map to one and only one belief set. This is in line with our assumption above that the joint epistemic state supervene on the epistemic state of the individuals involved.[10]

To account for the idea of expertise, we introduce the notion of an *expert rights system* which is an n-tuple $(R_1, ..., R_n)$, where each R_i is a (possibly empty) subset of X closed under negation. For each individual i, R_i contains the sentences concerning her area of expertise. With respect to an aggregation function F, call an individual i *decisive*

[10]Nevertheless, this means ruling out an important class of judgement aggregation procedures, namely so-called 'distance-based' ones which map profiles to belief sets that, in some sense, minimise the 'global' distance between the individual beliefs and the joint belief, since in general these are not resolute. These distanced-based approaches were inspired by belief-merging operators and introduced in the context of judgement aggregation by Pigozzi (2006), see Grossi and Pigozzi (2014, Section 4.3.3) for a useful presentation. They require a metric on the space of belief sets, and a way of combining the distance between individual beliefs into a distance between a profile and potential belief sets associated with that profile, and then for any given profile map it to the belief set which minimises this distance. There is (usually) no guarantee that this belief set be unique (without a tie-breaking rule), and for this reason, combined with the fact that they can only be used with a defined metric on belief sets and way of combining this into a distance between profiles and such sets, which may only be available in very specific situations, we set them aside here.

on a set of sentences $Z \subseteq X$ if and only if $F(A_1, \ldots, A_n) \cap Z = A_i \cap Z$, i.e. the sentences from Z that are contained in the joint report are all and only those believed by i.

Fragmented disciplines contain scientists who are experts over certain areas of research. The sentences (and their negations) from those areas go into the scientists' expert rights sets. Since they are the experts over these sentences, we model them as being decisive over whether or not the collaboration they are a part of, as a whole, should adopt them. We assume throughout that all rights sets in a rights system are disjoint. So for any sentence under consideration there corresponds at most one expert. If the same sentence appears in multiple rights sets, the impossibility result to come would be trivial.[11]

The idea that the beliefs of a collaboration should depend on the beliefs of the scientists within it is captured by the requirement of a judgement aggregation function. The below condition on the function demands that regardless of what the individual scientists believe (as long as they are individually consistent), they should be able to reach a collectively supported set of sentences following the agreed aggregation method:

Universal Domain: The domain of F is the set of all possible profiles of consistent judgement sets.

[11]Plausibly we can think about a collaboration involving experts as involving a 'representative expert' for each discipline. It doesn't matter for our current purposes how each representative expert has arrived at their belief set; the question is how to combine the beliefs of the representatives involved. However, it is worth nothing that there are ways of generalising the below to allow for expert rights of subgroups. We suppress them here for brevity, but see Dietrich and List (2008, pp. 64-65) for details.

The definition of Universal Domain is more general than the definition of Dietrich and List (2008, p. 63) (although see Dietrich and List, 2007b). They require the function F to be defined over the set of all possible profiles of consistent and *complete* judgement sets, where a judgement set A_i is complete with respect to an agenda X if and only if for any $\varphi \in X$, either φ or $\neg\varphi$ in A_i. Requiring completeness does not seem intuitive for scientists trying to arrive at a collective decision on what a collaboration tells us about nature. We needn't restrict ourselves to cases where every scientist involved in the aggregation procedure has an opinion on every sentence in the agenda. As we explain later in this section, this variation makes no difference to the impossibility result to come.

The below is a formally precise formulation of the suggestion that experts should determine what the collaboration 'believes' regarding their respective areas of expertise:

Minimal Expert Rights: There are at least two individuals i, j who are respectively decisive over non-empty R_i and R_j.

Finally, the fact that whenever everyone agrees on something it should be believed by the collaboration as a whole is given by:

Unanimity Principle: For any profile $(A_1, ..., A_n)$ in the domain of F and any sentence $\varphi \in X$, if $\varphi \in A_i$ for all individuals i, then $\varphi \in F(A_1, ..., A_n)$.

Whilst we take this principle to be well motivated in the context of scientific collaborations in general, it is worth noting that there may be certain specific cases where it can be questioned. In particular, there may be certain collaborative contexts which naturally call for the use of a 'premised-based' aggregation function. Such functions take advantage of the structure of specific agendas; namely those where the propositions in question can be partitioned into the ‘premises’ and the ‘conclusions’, where the former set of propositions are in some sense logically independent of one another (e.g. in the framework we’re working in they could form a disconnected sub-agenda), and are such that positions on the premises entail, according to the logic in question, positions on the conclusions. In such cases it might be plausible to adopt an aggregation function defined by aggregating profiles restricted to the premises, with the joint positions on the conclusions being determined by what’s entailed by the joint positions on the premises. Such an aggregation function can violate the Unanimity Principle in the sense that it can result in a joint position on a sentence from the set of conclusions (e.g. $\neg\varphi \in F(A_1, ..., A_n)$) even though every individual in the collaboration believe the opposite (e.g. $\varphi \in A_i$ for all i).

Whilst these sorts of aggregation functions may be appropriate in certain cases of collaboration, we’re putting them aside for our current purposes for two reasons. First, and most importantly, they require that the agendas in question have a relatively rich structure: it need be both that the set of premises are appropriately disconnected to ensure that the impossibility result we present below doesn’t reappear restricted to the premises sub-agenda, and that the premises (in combination with the logic) be such that the result of aggregating them be sufficient to entail positions on the conclusions (in particular, conclusions for which the joint report is expected to deliver a recommendation). It is far from clear whether or not all questions scientific collaborations are expected to answer have such a rich structure. Even in our toy example in the previous section, there was no obvious distinction between premises and conclusions, and we might expect that many cases of scientific collaboration are holistic in the sense that their aim is to knit together different scientific fields and areas of expertise, rather than merely trying to generate conclusions.

Second, there are worries with the very idea of giving up on the Unanimity Principle. As discussed previously we want to remain relatively non-committal about the relationship between the content of the joint report and the epistemic state of the collaboration, but it is worth noting that such functions may require that we attribute a doxastic state to the group (e.g. believing $\neg\varphi$) even though every member of the group believes the opposite (e.g. believe φ).[12] If this were to be the case there would be a distinctive gap between the doxastic state of the group and the states of its members. At the very least, more motivation would need to be given for the plausibility of such a gap. So, it suffices for our current purposes that there are at least some cases of scientific collaboration that don't have the premise-conclusion structure, or at least some cases where the Unanimity Principle be expected to hold.

In these sorts of cases, with this technical apparatus at hand, we can now present Dietrich and List's result establishing that a collaboration spanning fragmented scientific fields that attempts to reach a collective judgement via an aggregation function that respects the above conditions cannot protect itself against inconsistencies:

Theorem 1. *If the agenda is connected, there exists no aggregation function (generating consistent collective judgement sets) that satisfies Universal Domain, the Unanimity Principle, and Minimal Expert Rights.*

Proof. See Dietrich and List (2008, pp. 72-73). Notice that they restrict their focus to a

[12] Notice that in these cases every individual explicitly believes the negation of the group's position on the sentence in question, not merely that the group believe a proposition that no individual believes.

proper subset of the Universal Domain as we define it, i.e. the set of consistent and complete profiles. But if there is no function that satisfies Minimal Expert Rights and the Unanimity Principle over their domain, then, *a fortiori* there is no function that satisfies them over a superset of that domain (our Universal Domain). □

The toy example provided at the end of the previous section demonstrates the result. In terms of the formal model, the agenda under consideration is the following $\{p, \neg p, q, \neg q, p \rightarrow q, \neg(p \rightarrow q)\}$. Let individual 1 be the expert on p (carbon dioxide emissions). Let 2 be the expert on q (ocean temperature). Thus $R_1 = \{p, \neg p\}$ and $R_2 = \{q, \neg q\}$. Recall that 1 believes that carbon dioxide emissions are above the threshold, and that if they are, then oceans are becoming warmer and that 2 believes the conditional, but denies that oceans are becoming warmer, i.e. they believe the rise in oceans' temperature is temporary and not indicative of an upwards trend (maybe they don't trust historical measurements, maybe they believe increases in temperature are cyclical, etc.). Thus $A_1 = \{p, p \rightarrow q, q\}$ and $A_2 = \{\neg q, p \rightarrow q, \neg p\}$. By Universal Domain, F must be defined over the profile (A_1, A_2).[13] By the Unanimity Principle $p \rightarrow q \in F(A_1, A_2)$. By Minimal Expert Rights $p, \neg q \in F(A_1, A_2)$. Thus, $F(A_1, A_2)$ is inconsistent.[14]

[13]Note that this profile is also 'unidimensionally aligned' so would also be included in a domain restricted to unidimensionally aligned profiles, a common domain restriction in the judgement aggregation framework. Such a restriction avoids some impossibility results (see List (2003) and Grossi and Pigozzi (2014, Section 4.1.1)), but not the one we are considering here.

[14]Note that in this example both 1 and 2 have *complete* judgement sets over the items in the agenda. One might worry that the inconsistencies that arise require completeness of

4 A formal model of intellectual modesty in scientific collaborations

The impossibility result of the previous section shows that if the members of a collaborations are allowed to be decisive over the propositions in their area of expertise, the collaboration as a whole might be forced into an inconsistency. In response to this, we believe scientists should relinquish their special claims over the propositions in their area of expertise and contribute them to the aggregation on a par with everyone else. We take this strategy to be one of intellectual modesty. In this section we offer a formal account of intellectual modesty in the judgement aggregation framework articulated in Section 3, while in the next we argue for its plausibility in the contexts under consideration.

Inspired by Gibbard's (1974) response to Sen's paradox, we believe an appropriate technical notion that would capture the move towards intellectual modesty when confronted with an inconsistency can be given by allowing expertise to be alienable. It can be introduced into the formal model we are working with here as follows. First,

---

each agents' judgement set, and thus Theorem 1 can be avoided if we restrict the domain to profiles of incomplete judgement sets, and that it's these scenarios that are relevant in the contexts under consideration, viz. scientific collaborations (where scientists typically won't hold attitudes towards every proposition under consideration). This is not the case. Consider a super-set of the above agenda which includes the proposition r, and all of its logical relations with p and q. Then, let 1 and 2's judgement sets be as before (and therefore be incomplete with respect to the enriched agenda). Clearly the contradiction is still generated.

define an individual i's *waiver set* $W_i \subseteq R_i$ as follows:

1. Define a set Ψ (relative to a given profile and rights system):

$$\Psi =_{df} \{\psi \in X \colon [\forall k \in \{1, \ldots, n\} \colon \psi \in A_k] \vee [\exists k \in \{1, \ldots, n\} \colon \psi \in (A_k \cap R_k)]\}$$

2. For any sentence $\varphi \in R_i$:

$$\varphi \in W_i \Leftrightarrow \exists \Psi' \subseteq \Psi \colon [\Psi' \text{ is consistent }] \wedge [\Psi' \cup \{\varphi\} \text{ is inconsistent}]$$

Ψ is the set of sentences endorsed by the community in the sense that they would be guaranteed to go into the collective judgement set by unanimity or minimal expert rights, should they hold. The waiver set of an individual i, W_i, is the set of claims that if i were to exercise her right over them, they would generate collective inconsistency when combined with a consistent subset of Ψ (if everyone else also exercised their rights over their fields of expertise).

This then provides the following alternative to Minimal Expert Rights:

Alienable Expert Rights: For every profile, and every individual i, if i is an expert over φ, i accepts φ, and φ is not waived, then φ is included in the collaboration's collective belief set, i.e. if $\varphi \in (A_i \cap R_i \setminus W_i)$, then $\varphi \in F(A_1, \ldots, A_n)$.

Notice that this condition is analogous to what Dietrich and List (2008, p. 64) call *positive decisiveness* which requires that an individual i is positively decisive over their rights set if and only if $(A_i \cap R_i) \subseteq F(A_1, \ldots, A_n) \cap R_i$, rather than *negative decisiveness*, which requires $F(A_1, \ldots, A_n) \cap R_i \subseteq (A_i \cap R_i)$.[15] In the context under consideration we take the positive, rather than negative, aspect of this notion to capture the relevant sense of expertise: if an expert believes a sentence then, assuming it is not waived, it should be part of the collaboration's (collective) judgement set. To briefly motivate this, consider the case in which an expert avoids taking a position on a sentence from their area of expertise (i.e. neither believes it nor its negation) – the sentence could correspond to an 'open problem' in their field – then through negative decisiveness that sentence could not become part of the collaboration's collective judgement. But it seems unintuitive to prevent scientists working in related fields from settling the question. However, the below result continues to hold when the condition is strengthened to demand that if the group accepts φ, and φ is in R_k for some k, then A_k (the analogue of negative decisiveness).

With this in place, we provide the following possibility result:

Theorem 2. *For any connected agenda and any rights system, there exists an aggregation function (generating consistent collective judgement sets) that satisfies Universal Domain, the Unanimity Principle, and Alienable Expert Rights.*

Proof. The following function, F, respects Universal Domain, the Unanimity Principle, and Alienable Expert Rights by construction:

[15]We say 'analogous to' as the introduction of the waiver condition entails that we cannot use decisiveness in the sense of Dietrich and List (2008, pp. 63-64), since in Alienable Expert Rights, what experts are 'decisive over' varies from profile to profile. Also notice that Theorem 1 holds when Minimal Expert Rights is weakened to positively Minimal Expert Rights (Dietrich and List, 2008, p. 64).

$$F(A_1, \dots A_n) = \bigcap_i^n A_i \cup \bigcup_i^n (A_i \cap R_i \setminus W_i)$$

It remains to demonstrate that it is guaranteed to generate a consistent collective judgement set. We show this by induction on its subsets, for some arbitrary profile.[16] To begin with, $\bigcap_i^n A_i$ is consistent by the consistency of every individual judgement set. Suppose then Ω is a subset of $\bigcup_i^n (A_i \cap R_i \setminus W_i)$ such that $\bigcap_i^n A_i \cup \Omega$ is consistent. The inductive step is that if $\bigcap_i^n A_i \cup \Omega$ is consistent then so is $\bigcap_i^n A_i \cup \Omega \cup \{\varphi\}$, for any $\varphi \in \bigcup_i^n (A_i \cap R_i \setminus W_i)$. Assume towards a contradiction that this is not the case, i.e. $\bigcap_i^n A_i \cup \Omega \cup \{\varphi\}$ is inconsistent, for some φ (1). Notice that $\bigcap_i^n A_i \cup \Omega$ is a consistent subset of Ψ (2). Since $\varphi \in \bigcup_i^n (A_i \cap R_i \setminus W_i)$, then there must exist a unique k (by the disjointness of the rights system), such that $\varphi \in (A_k \cap R_k \setminus W_k)$. But $(A_k \cap R_k \setminus W_k) \subseteq (A_k \cap R_k)$ and hence $\varphi \in (A_k \cap R_k)$ (3). From (1), (2), and (3): $\varphi \in W_k$ and $\varphi \notin \bigcup_i^n (A_i \cap R_i \setminus W_i)$. This contradicts our assumption and since our choice of φ was arbitrary, $\bigcap_i^n A_i \cup \Omega \cup \{\varphi\}$ is consistent for any $\varphi \in \bigcup_i^n (A_i \cap R_i \setminus W_i)$.[17] This concludes the proof that F always generates a consistent collective judgement set. □

In a collaboration spanning fragmented disciplines where the expert rights system and scientists' beliefs are such that respecting these rights (and the Unanimity Principle) will yield an inconsistent result, the collaboration has gone wrong somewhere.[18] Even if

[16]Recall we restrict our focus to monotonic logics throughout.

[17]Recall our initial assumption that the agenda is finite. In consequence, the set $\bigcup_i^n (A_i \cap R_i \setminus W_i)$ will also be finite. If we were to relax our assumption, the proof would follow along the same lines, but we would have to make an additional assumption that the underlying logic is compact.

[18]We take this to be uncontroversial and to hold even if more than one doxastic attitude

the scientists' expert rights sets are disjoint, if there are enough logical connections between the sentences under consideration, then an inconsistency may be generated. Alienable Expert Rights suggests that whenever a collaboration finds itself in such a situation, every expert should waive their expertise over claims which contribute to the inconsistency (which needn't be all of them). And in doing so they contribute to what goes into the collective judgement on a different footing. Assuming the Unanimity Principle, the expertise of some individuals cannot be respected whilst respecting the expertise of others. So, if scientists were to be intellectually modest, the role of this expertise should be reconsidered.

Notice that the notion of Alienable Expert Rights is quite demanding: it requires that *every* scientist who is an expert over *any* sentence contributing to the inconsistency waive their expertise over those sentences. There are pros and cons to defining Alienable Expert Rights in this way. On the positive side, it can be used in cases where there is no reason to respect the expertise of one expert over another. It is egalitarian; all experts are treated on a par, and in cases of inconsistency they *all* need to waive their expertise.[19]

---

may be justified by an existing body of evidence (see Kopec and Titelbaum, 2016, for a discussion of the Uniqueness Thesis).

[19]In this sense the intuitive motivation for defining alienable expertise in this way reflects the egalitarian integrity constraint (called IC4) in the belief-merging literature (see e.g. Konieczny and Pérez, 2011; Pigozzi, 2016). This constraint requires that a belief merging operator be such that the result of merging a profile into a joint/merged belief set be consistent with the belief set of some individual in that profile if and only if it's consistent with the others. The fusion operators which respect it don't prioritise individuals involved,

On the negative side, in certain cases some scientists may have a legitimate claim to a more secure evidential basis than others. Different sub-disciplines may place different demands on whether an individual should include a sentence in their belief set, and this may be salient when it comes to deciding who should waive their expertise, i.e. perhaps individuals with less secure evidence for their beliefs should waive before individuals with more secure evidence (cf. Martini and Sprenger, 2018). Since our model does not include a parameter measuring strength of evidence, it is insensitive to such concerns; we simply assume that all scientists are equally well justified in holding their expert beliefs. And in those cases, our approach to Alienable Expert Rights would require all experts to waive their expertise, even though it may be sufficient for a single expert (perhaps one who comes from a sub-discipline with a low evidential standard) to do so. Which some may see as an unsatisfactory result.

This point is well taken. But it is worth pointing out that there is no reason to suppose that the model could not be further developed to take this into account, by introducing an order of priority in which experts waive their expertise (and then redefining Alienable Expert Rights to ensure that in the problematic cases just enough experts waive their expertise to guarantee a consistent joint belief).

Moreover, we are not suggesting that, F, as defined above, be the method scientists should use when forming scientific consensus. It is plausible that the collective belief sets it delivers will be rather limited in many cases. For any sentence φ that is not

---

just as our notion of alienable expertise doesn't prioritise one expert at the expense of another, since either their expertise doesn't contribute to a contradiction, in which case their expertise is respected, or their expertise does contribute to a contradiction, in which case all of them need to waive their expert status over the sentences involved.

universally agreed upon, or in someone's area of expertise, φ won't be in the value of F. And we can imagine cases involving a group of extraordinarily timid experts who do not take any position on the sentences they are experts over (or alternatively, where they are required to waive their expertise over their entire rights set), and moreover do not unanimously agree on any sentence in the agenda. In such cases, F will deliver the empty set. And this would indeed be too restrictive. Our claim is that any acceptable aggregation function should be such that its value for any profile be a superset of the value of F on that profile.

Another way of putting it is that we take Universal Domain, Alienable Expert Rights, and the Unanimity Principle to be necessary conditions on any acceptable aggregation function, but they needn't be sufficient. An interesting question then, is what constraints the necessary conditions put on expansions of F.[20]

To the best of our knowledge there has been no investigation into the question of what constraints Universal Domain and Unanimity put on judgement aggregation functions. However, there are some results from the belief-merging literature that can be brought to bear on this question. Define a judgement aggregation function to be a *Generalised Varying Dictatorship* if for any profile $(A_1, ..., A_n)$ in its domain, $F(A_1, ..., A_n) = A_i$ for some (not necessarily fixed across profiles) individual i.[21] Then,

[20]Notice that if Alienable Expert Rights is strengthened to include the analogue of negative decisiveness as discussed above, then this will place an additional constraint on any expansion of F.

[21]This corresponds to the notion of a 'generalised dictatorship' in Grandi and Endriss (2013). We have included the modifier 'varying' in its name to additionally stress that the individual the group agrees with can vary across the domain. A varying dictatorship is

we can prove:

Theorem 3. *If the logic is propositional, the agenda X is closed under disjunction, and individual and joint belief sets are complete (in all profiles in the domain), a judgement aggregation function (that is guaranteed to deliver a consistent and complete collective judgement sets) respects the Unanimity Principle if and only if it is a Generalised Varying Dictatorship.*

Proof. This proof follows Grandi and Endriss's strategy for proving a more general theorem (2013, Theorem 16). There it is made in the belief-merging framework. Here we rephrase it in the context of judgement aggregation. Assume that the logic is propositional, that X is closed under disjunction, and that the individual and joint belief sets are complete in all profiles in the domain. That a Generalised Varying Dictatorship respects the Unanimity Principle is obvious: if $F(A_1, ..., A_n) = A_i$, any sentence believed by all is believed by A_i and is thus included in $F(A_1, ..., A_n)$. For the left-to-right, assume towards contradiction that F respects the Unanimity Principle and is not a Generalised Varying Dictatorship. By the latter assumption there exists a profile $(A_1, ...A_n)$ such that $F(A_1, ..., A_n) \neq A_i$ for any A_i. So for all A_i, there is a (not necessarily the same across individuals) sentence $\varphi_i \in X$ such that either $\varphi_i \in A_i$ and $\varphi_i \notin F(A_1, ..., A_n)$ or $\varphi_i \notin A_i$ and $\varphi_i \in F(A_1, ..., A_n)$. Define $\pm\varphi_i$ as φ_i if $\varphi_i \in A_i$, and $\neg\varphi_i$ if $\varphi_i \notin A_i$. Either way, for all i, since the individual belief sets are complete, $\pm\varphi_i \in A_i$, and since the joint belief set is complete $\neg \pm \varphi_i \in F(A_1, ..., A_n)$. Since X is closed under disjunction (and assuming a finite number of individuals), it includes the disjunction of the $\pm\varphi_i$s, i.e. $\bigvee \pm\varphi_i \in X$. Since each individual belief set A_i is complete and consistent, and $\pm\varphi_i \in A_i$,

not the same as a dictatorship and shouldn't be taken to have the same connotations.

$\bigvee \pm\varphi_i \in A_i$, for all i. So by the Unanimity Principle $\bigvee \pm\varphi_i \in F(A_1, ..., A_n)$. But $\neg \pm \varphi_i \in F(A_1, ..., A_n)$, for all i. So $F(A_1, ..., A_n)$ is inconsistent. $\square$

This proof relies heavily on the assumption that the individual and joint belief sets be complete, which as we discussed above, should not be expected in many cases of scientific collaboration. In contexts where incompleteness is expected, the relationship between the Unanimity Principle and Generalised Varying Dictatorship, or more specifically, whether or not it rules out profiles where $F(A_1, ..., A_n) \nsubseteq A_i$ for any i, or $A_i \nsubseteq F(A_1, ..., A_n)$ for any i, need to be investigated.

Exploring additional conditions strikes us as a potentially fruitful avenue for future research, but in this paper we are only interested in providing a possibility result: there are functions defined on the universal domain which satisfy the Unanimity Principle and Alienable Expert Rights, how they can be expanded goes beyond our current purposes.

5 Intellectual modesty and indirect peer disagreement

By replacing the strict notion of expertise modelled by Minimal Expert Rights, with the weaker notion corresponding to Alienable Expert Rights, the impossibility result can be avoided. The collaboration should allow experts to submit their beliefs with respect to the sentences that led to the contradiction, but the individual experts should no longer be given the ability to ensure they are each included in the joint belief. It remains then, to demonstrate that the notion of alienable expertise is defensible, and to show that it is to be preferred to the other possibility results present in the literature. These are our tasks in this section.

In a way, the problem experts encounter in the kinds of cases that demonstrate

Theorem 1 is akin to the problem discussed in the peer disagreement literature (see, for instance, Feldman, 2006; Christensen and Lackey, 2013). In that context, two epistemic peers with respect to the same proposition φ find that they hold inconsistent beliefs about φ. The question there is how to rationally react when your peer disagrees with you? (for an overview see Frances and Matheson, 2019, Section 5).

Things are slightly different in our context. Recall the simple example of Section 3 in which individual 1 was an expert on p and her judgement set included the propositions p, $p \to q$ and q and individual 2 was an expert on q and her judgement set included $\neg q$, $p \to q$ and $\neg p$. Individual 1 is in an epistemically superior condition with respect to proposition p, individual 2 is an epistemically superior position with respect to proposition q and they are both equally well positioned epistemically with respect to $p \to q$. Their joint belief set was $\{p, p \to q, \neg q\}$ and was hence inconsistent. But the two scientists in that example are not epistemic peers with respect to p or q (and indeed don't disagree with respect to $p \to q$). So at first sight the disagreement we are concerned with in this paper does not seem to be an instance of peer disagreement.[22]

However, since it's p and q, combined with some unanimously believed propositions, which constitute the disagreement that arises in the form of the contradictory joint belief

[22]Another way of looking at this case, however, would be as one of *superficial* agreement on the proposition $p \to q$ masking a *deeper* case of peer disagreement. While the scientists involved are in agreement that $p \to q$, this is only because one of them has the evidence p while the other the evidence $\neg q$. So in effect, the apparent agreement would vanish, on pain of inconsistency, if they were to defer to each others' expertise. For a discussion of superficial agreement and deep disagreement see Rowbottom (2018). For why deferring is not the appropriate rational strategy in scientific collaborations, see below.

set, the case that demonstrates Theorem 1 can instead be viewed as a novel case of *indirect* peer disagreement. While the scientists involved are not epistemic peers in the sense of being equally familiar with the evidence pertaining to the propositions that lead to the contradiction (Kelly, 2005), there is a more general sense in which they are concerning the agenda as the whole. The scientists are peers in terms of their expertise with regards to the propositions in their area of specialization which lead to the contradictory joint belief set.

The resolution of indirect cases of peer disagreement, nevertheless, may appear *prima facie* simpler than what is usually discussed in the epistemology of disagreement literature. Since the scientists involved are each others' epistemic superiors with respect to the propositions in their area of expertise, one simple way out of this problem would be for them to defer to one another on those propositions (see, for instance Frances, 2013, for a discussion of the conditions under which this is the only rational strategy).

This way of thinking connects to some of the possibility results discussed by Dietrich and List (2008). They show that by restricting the domain to profiles that contain at least one *deferring* individual (where i is deferring in a profile $(A_1, ..., A_n)$ if and only if $A_i \cap R_j = A_j \cap R_j$ for every $j \neq i$), or at least one *agnostic* individual (where i is agnostic in a profile $(A_1, ..., A_n)$ if and only if A_i is consistent with every set of the form $B_1 \cup ... \cup B_{i-1} \cup B_{i+1} \cup ... B_n$ for each $j \neq i, B_j \subseteq R_j$)[23], then a possibility result can be generated by setting the value of the aggregation function to the belief set of the deferring or agnostic individual (in the latter case combined with the other experts'

[23]Notice that in such cases an agnostic individual may have an incomplete judgement set, but also notice that incompleteness at a profile does not entail agnosticism at that profile.

beliefs regarding their areas of expertise).[24] These results require collaborations where at least one scientist simply will not disagree with what the experts say about their respective areas of expertise. Such a scientist has to either defer to every other expert, or withhold judgement over any sentence in someone else's expert rights set.[25] In a way this amounts to an individual solution to a collective problem: one or more individuals would have to revise their beliefs in the interest of group consistency.

Now, such a strategy may be rational for a scientist if they believe they are significantly epistemically inferior to their colleagues at least with respect to the propositions in their areas of expertise. Nevertheless, scientists working in a collaboration towards the same epistemic goal do hold justified beliefs over areas outside their own expertise, and those beliefs are such that they should be respected by the collaboration as whole. Their beliefs, at least about the propositions the collaboration as a whole may take an interest in, are connected. Engineers working on interferometers

[24]They also investigate a possibility result regarding severely restricted agendas and expert rights systems, but assuming the scientists are aiming to achieve a consensus which could contain logically connected sentences, the type of result List and Dietrich have in mind is implausible in this context.

[25]A further problem is that the way the possibility results are generated does suggest that scientists might be motivated to defer, or remain agnostic, since in doing so their belief set determines the value of the aggregation function (Dietrich and List, 2008, p. 74). Although they do not act as a dictator in the Arrovian sense (the same named individual determines the value of the function for any profile), they do determine whether or not any sentence not in anyone's rights set goes into the collective judgement set. And this is not a desirable result.

and astrophysicists studying black holes have in some sense very separate areas of expertise. That being said, within LIGO-Virgo, in the attempt to build interferometers capable of detecting gravitational waves generated by the merging of a binary black hole system scientists had to blur the boundaries of their respective fields of expertise. In this sense, we might think of them as 'quasi-peers': sufficiently well-acquainted with the area in question, and certainly acquainted enough to be considered valuable interlocutors on the issues, although not enough to be considered 'full' experts in their own right.

Commenting on the experience of sociologists of science, Collins and Evans (2002, p. 254) distinguishes between three levels of 'expertise': no expertise, contributory expertise and interactional expertise. The first notion is self-explanatory. To have contributory expertise is to have "enough expertise to contribute to the science of the field" (ibid.), whereas to have interactional expertise is to have "enough expertise to interact interestingly with participants" (ibid.) who have contributory expertise. Our claim then, is that in such collaborations, scientists who hold (justified) beliefs on areas in others' areas of expertise, may nevertheless be considered interactional (but not contributory) experts.

With this distinction at hand, the strongest version of the counterargument under consideration that individual scientists should defer/remain agnostic on propositions in others' areas of expertise, is that scientists working in a collaboration do not have the interactional expertise to meaningfully engage with the reasons colleagues provide in support of their position on a scientific matter. This, however, is implausible.

Returning to the case of the LIGO-Virgo collaboration, Collins argues that having interactional expertise in all the fields of study represented in the collaboration is required for a scientist to make a contribution to the epistemic goals of the collaboration in line with their expertise:

> Multiple authorship does reflect the complex division of labor in this field, but it also reflects a shared understanding, or potential shared understanding, of every aspect of the work. Indeed, without this possibility it would be hard to understand how the division of labor or the management of the whole project would work. Therefore, in principle, each of the ... authors, though each contributes to the project in a different way, can still be a full author of the paper. (Collins, 2017, p. 368)

This is what provides scientists within a collaboration with the epistemic legitimacy to weigh and perhaps reject the reasons that colleagues provide in support of their expert opinions. The same argument as to why they can collaborate on interdisciplinary projects works as an argument against them deferring to each others' expert claims. Therefore, solutions to the indirect peer disagreement problem, such as those explored by Dietrich and List (2008), are not available to scientists working in a large collaboration.

In particular, in such contexts experts should have to *justify* why what they are contributing should be accepted by the collaboration. The idea that experts should be able to justify their expertise is not new. Martini (2014, pp. 5-7) puts forward as a principle of expertise that "[e]xperts should support their theses with arguments and evidence" (see also Goldman, 2001, pp. 93-97). Experts have the opportunity to present their reasons for holding a certain view on a scientific matter that falls within their area of expertise and persuade the rest of the collaboration. Their beliefs should win the day in a collaboration because their reasons are persuasive enough, not simply because they are experts on the subject matter. If they don't, it means the reasons were not sufficiently persuasive and hence there are legitimate reasons to accept a different interpretation of the evidence. Therefore, experts should all agree to waive their expert rights over the sentences that led to the inconsistency, and agree to perform the aggregation as if none of them were experts over these sentences.[26].

6 Conclusion

Despite its apparent desirability, the fragmentation and specialisation of academic disciplines poses a problem. How are we to determine the state of a scientific collaboration as a whole? Under certain plausible assumptions, Dietrich and List's impossibility result shows that this cannot be answered without the threat of inconsistency. We proved that this can be avoided by replacing expert rights with alienable expert rights, when an expert right is alienable when it conflicts with expert beliefs and the beliefs unanimously accepted. We argued that this provides a more plausible way of avoiding inconsistency in scientific collaborations than the (anti-egalitarian) idea of venerating expertise at the expense of respecting the beliefs of the individuals involved.

This result has implications for the possibility of consistency of scientific collaborations, and consequently for their status as group agents. In 2016, the LIGO-Virgo Collaboration (LVC) announced the first direct detection of a gravitational wave. Their discovery paper, published in *Physics Review Letters*, included the following statement:

[26]This does not preclude retaining expertise over some elements in the agenda that did not contribute to the inconsistency in the collective belief set

> A century after the fundamental predictions of Einstein and Schwarzschild, *we* report the first direct detection of gravitational waves and the first direct observation of a binary black hole system merging to form a single black hole. (Abbott et al., 2016, p. 061102-1, our emphasis)

But in what sense do 1000+ scientists spanning multiple countries, languages, institutions, and fields of expertise speak with one voice and act with one mind? Following List and Pettit (2005), we can assume a minimal condition for a scientific collaboration to act as a group agent is to support consistent views. This paper suggests that in order to achieve this minimum threshold scientists ought to be willing to waive their right to dictate over their respective fields of expertise and contribute their expert beliefs on a par with everyone else's.

References

Abbott, B. P. et al. (2016, Feb). Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett. 116*, 061102.

Bird, A. (2010). Social knowing: The social sense of scientific knowledge. *Philosophical Perspectives 24*(1), 23–56.

Bird, A. (2014). When is there a group that knows? distributed cognition, scientific knowledge, and the social epistemic subject. In *Essays in collective epistemology*, pp. 42–63. Oxford University Press.

Bright, L. K., H. Dang, and R. Heesen (2018). A role for judgment aggregation in coauthoring scientific papers. *Erkenntnis 83*(2), 231–252.

Casadevall, A. and F. C. Fang (2014). Specialized science. *Infection and Immunity 82*(4), 1355–1360.

Christensen, D. and J. Lackey (2013). *The epistemology of disagreement: New essays*. OUP Oxford.

Coccia, M. and L. Wang (2016). Evolution and convergence of the patterns of international scientific collaboration. *Proceedings of the National Academy of Sciences 113*(8), 2057–2061.

Collins, H. (2017). *Gravity's Kiss: The Detection of Gravitational Waves*. MIT Press.

Collins, H. and R. Evans (2002). The third wave of science studies. *Social Studies of Science 32*(2), 235–296.

Constantian, M. B. (1999). The gordian knot of multiple authorship. *Plastic and Reconstructive Surgery 103*, 2064–2066.

Cronin, B. (2001). Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology 52*(7), 558–569.

Dietrich, F. (2006). A generalised model of judgment aggregation. *Social Choice and Welfare 28*(4), 529–565.

Dietrich, F. and C. List (2007a). Arrow's theorem in judgment aggregation. *Social Choice and Welfare 29*(1), 19–33.

Dietrich, F. and C. List (2007b). Working paper: Judgment aggregation with consistency alone. Maastricht University.

Dietrich, F. and C. List (2008). A liberal paradox for judgment aggregation. *Social Choice and Welfare 31*(1), 59–78.

Dietrich, F. and C. List (2013). Propositionwise judgment aggregation: the general case. *Social Choice and Welfare 40*(4), 1067–1095.

Dizadji-Bahmani, F., R. Frigg, and S. Hartmann (2010). Who's afraid of Nagelian reduction? *Erkenntnis 73*(3), 393–412.

Feldman, R. (2006). Reasonable religious disagreements. In L. Antony (Ed.), *Philosophers Without Gods: Meditations on Atheism and the Secular Life*, pp. 194–214. Oxford University Press.

Frances, B. (2013). Philosophical renegades. In J. Lackey and D. Christensen (Eds.), *The Epistemology of Disagreement: New Essays*, pp. 121–166. Oxford University Press.

Frances, B. and J. Matheson (2019). Disagreement. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2019 ed.). Metaphysics Research Lab, Stanford University.

Galison, P. (2003). The collective author. In P. Galison and M. Biagioli (Eds.), *Scientific Authorship: Credit and Intellectual Property in Science*, pp. 325–353. Routledge.

Gibbard, A. (1974). A Pareto-consistent libertarian claim. *Journal of Economic Theory 7*(4), 388–410.

Gilbert, M. (2000). *Sociality and Responsibility: New Essays in Plural Subject Theory.* Rowman & Littlefield Publishers.

Goldman, A. I. (2001). Experts: Which ones should you trust? *Philosophy and Phenomenological Research 63*(1), 85–110.

Grandi, U. and U. Endriss (2013). Lifting integrity constraints in binary aggregation. *Artificial Intelligence 199-200*, 45 – 66.

Grossi, D. and G. Pigozzi (2014). Judgment aggregation: A primer. *Synthesis Lectures on Artificial Intelligence and Machine Learning 8*(2), 1–151.

Huebner, B., R. Kukla, and E. Winsberg (2018). Making an author in radically collaborative research. In T. Boyer-Kassem, C. Mayo-Wilson, and M. Weisberg (Eds.), *Scientific Collaboration and Collective Knowledge*, pp. 95–116. Oxford University Press.

Jones, B. F., S. Wuchty, and B. Uzzi (2008). Multi-university research teams: Shifting impact, geography, and stratification in science. *Science 322*(5905), 1259–1262.

Kelly, T. (2005). The epistemic significance of disagreement. In J. Hawthorne and T. Gendler (Eds.), *Oxford Studies in Epistemology, Volume 1*, pp. 167–196. Oxford University Press.

Konieczny, S. and R. P. Pérez (2011). Logic based merging. *Journal of Philosophical Logic 40*(2), 239–270.

Kopec, M. and M. G. Titelbaum (2016). The uniqueness thesis. *Philosophy Compass 11*(4), 189–200.

List, C. (2003). A possibility theorem on aggregation over multiple interconnected propositions. *Mathematical Social Sciences 45*(1), 1 – 13.

List, C. and P. Pettit (2002). Aggregating sets of judgments: An impossibility result. *Economics and Philosophy 18*(1), 89–110.

List, C. and P. Pettit (2005). On the many as one: A reply to Kornhauser and Sager. *Philosophy and Public Affairs 33*(4), 377–390.

Martini, C. (2014). Experts in science: a view from the trenches. *Synthese 191*(1), 3–15.

Martini, C. and J. Sprenger (2018). Opinion aggregation and individual expertise. In T. Boyer-Kassem, C. Mayo-Wilson, and M. Weisberg (Eds.), *Scientific Collaboration and Collective Knowledge*, pp. 180–201. Oxford University Press.

McDonald, K. A. (1995). Too many co-authors? *Chronicle of Higher Education April*(28), A35–A36.

Okruhlik, K. (1994). Gender and the biological sciences. *Canadian Journal of Philosophy 24*(sup1), 21–42.

Pigozzi, G. (2006). Belief merging and the discursive dilemma: an argument-based account to paradoxes of judgment aggregation. *Synthese 152*(2), 285–298.

Pigozzi, G. (2016). Belief merging and judgment aggregation. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 ed.). Metaphysics Research Lab, Stanford University.

Price, D. J. d. S. (1963). *Little Science, Big science*. Columbia University Press.

Rowbottom, D. P. (2018). What is (dis) agreement? *Philosophy and Phenomenological Research 97*(1), 223–236.

Sen, A. (1970). The impossibility of a Paretian liberal. *Journal of Political Economy 78*(1), 152–157.

Wray, B. K. (2001, Dec). Collective belief and acceptance. *Synthese 129*(3), 319–333.

Wray, B. K. (2002). The epistemic significance of collaborative research. *Philosophy of Science 69*(1), 150–168.

Wray, B. K. (2018). The impact of collaboration on the epistemic cultures of science. In T. Boyer-Kassem, C. Mayo-Wilson, and M. Weisberg (Eds.), *Scientific Collaboration and Collective Knowledge*, pp. 117–134. Oxford University Press.

Wuchty, S., B. F. Jones, and B. Uzzi (2007). The increasing dominance of teams in production of knowledge. *Science 316*(5827), 1036–1039.