Group Normic Reliabilism

Andrew Peet

Abstract

Our theorizing about collective and individual justification should be broadly continuous. As a result, we can learn a great deal about justification in general by focusing on the epistemic states of collectives: we can test our best theories of justification by seeing how they hold up as theories of collective justification. This paper discusses and generalizes a number of recalcitrant problems from the literature on collective justification and suggests that the difficulties we face in resolving these problems are damning not only for specific theories of collective justification they concern, but also the more general pictures of epistemic justification from which they derive. After all, if these general theories of justification were viable then they should extend straightforwardly to the justification of collective beliefs. However, all is not lost. I argue for the following theory of justification:

Normic Reliabilism A belief is justified if and only if it is true at all the most normal worlds at which it is held in the same way.

Not only is Normic Reliabilism a promising approach to individual justification, it also resolves the recalcitrant problems I discuss that are faced the majority of approaches to collective justification.

1. Introduction

Our epistemic theorizing tends to focus on individuals. The majority of our theories of justification were originally developed with the epistemic states of organisms like us in mind. However, individuals are not the only fitting objects of epistemic appraisal. We regularly epistemically evaluate the beliefs of groups. For example, the World Health Organization, Qanon fanatics, or the Parliamentary Labour Party.

We can learn a great deal about justification by focusing on the epistemic states of collectives. This is because our theorizing about collective and individual justification should be broadly continuous: our theories of justification in general must be satisfactory both as applied to organisms and collectives (at least, if we take group belief ascriptions seriously, as I will assume we should). So, we can test our best theories of justification by seeing how they hold up as theories of collective justification.

In this paper I will explore the following challenges:
1. The nature of collective belief, like that of individual belief, is highly controversial. Thus, given our current knowledge, any theory of justification that is tied to a particular conception of belief will, ceterus paribus, have a lower likelihood of being true than any theory that is neutral with respect to the nature of belief. Unfortunately, it has proved difficult to produce a viable theory of collective justification that does not place controversial demands on the nature of group belief.

2. Reliabilism is one of the most promising approaches to individual justification. However, as Lackey (2016) has illustrated, standard reliabilist approaches to collective justification divorce collective justification from the justification of individual group members. Very large groups of unreliable belief formers can collectively attain an extremely high degree of reliability and, thus, justification. This is a conclusion that, I shall argue, we should be keen to avoid.

3. Justified beliefs cannot be intentionally formed or brought about (by the holder of the belief) for purely practical purposes. Any view of justification that allows for justified beliefs to be brought about this way must be rejected. Yet, the majority of approaches to group justification allow for such beliefs to be justified (including approaches explicitly designed to avoid this result).

The majority of approaches to collective justification fail with respect to one or more of these challenges. Moreover, insofar as we should strive for continuity between our individual and collective epistemic theorizing, these problems are damning not only for specific theories of collective justification they concern, but also the more general pictures of epistemic justification from which they derive.

Yet, these challenges can be resolved by the following view of justification:

Normic Reliabilism A belief is justified if and only if it is true at all the most normal worlds at which it is held in the same way.

This is a form of normal worlds reliabilism, and retains the virtues of similar approaches (see Goldman (1986), Leplin (2006), Smith (2010, 2016, 2018), and Graham (2016, 2017)). It also dovetails with knowledge first approaches to justification. Moreover, it places almost no demands on the nature of
collective belief. It explains why large groups of unreliable belief formers do not typically form justified beliefs. And, it explains why groups cannot attain justified beliefs at will. Given its virtues as a theory of both individual and collective justification, the case for normic reliability in general is very strong.

I will begin by introducing the guiding methodological principle of this paper: that our theorizing about collective and individual epistemology should be continuous. I will then introduce normic reliability and explain how it satisfies the continuity desideratum. Next, I will turn to the three challenges outlined above. Finally, I close by responding to two objections to normic reliability.

2. The Principle of Continuity

When theorizing about group justification our theories should be continuous with our best theories of individual justification. Such continuity makes for simpler more elegant theories. Moreover, if we fail to maintain such continuity we are at risk of changing the subject when we turn to collective epistemology. Groups and individuals are importantly different. It is not clear from the outset that the types of epistemic appraisal fitting for groups and individuals will line up perfectly. Groups may admit of forms of epistemic appraisal that simply do not apply to individuals. For example, some groups, unlike organisms, have a degree of intentional control over their own constitution and design. Such groups may be epistemically criticizable in light of this. It is not clear that there is a direct analogue of such epistemic appraisal for individual organisms. Of course, we can point out that types of organism (e.g. humans) are built it sub-optimal ways. What we cannot do is point to a particular individual and criticize that very individual for its own design. This is something that, plausibly at least, we can do for some groups. Moreover, as I’ll suggest in §6, it may be more epistemically fitting to assess some groups as would would assess instruments, rather than agents. If we conflate prominent forms of positive (or negative) group epistemic appraisal with justification (or lack thereof), without any attempt to maintain continuity, we risk changing the subject and ceasing to talk about justification at all.

Moreover, when we theorize about justification we take ourselves to be theorizing about justification in general, not justification for a certain type of entity. We see justification as a unified property that does explanatory work. If our theories of justification are not continuous then this is called into question. It seems clear that individual justification matters. But if individual justification consists in the possession of property F, and group justification in the possession of property G, then it may be the possession of
F, rather than the possession of justification in general, that matters epistemically. In that case, we should ask not what it is for groups to be justified, but rather what it is for groups to possess property F (or approximate possession of F).

Gilbert and Plichman (2014) argue that our theories of collective phenomena should not be overly constrained by our corresponding theorizing in the individual realm:

‘In contemporary epistemology inquiries into the nature of belief and so on have started with the individual case. The results of these inquiries may well help to throw light on the nature of the collective case. Nonetheless caution is required with respect to the use of results deriving from the individual case in approaching the collective case. There may be significant disanalogies between these cases, despite their having some features in common. Careful, independent investigation of the collective case—as of the individual case—is required in order properly to understand it. Quite possibly, such careful investigation of the collective case will help to throw light on the individual case as well.’


I am largely in agreement with this sentiment. We cannot focus only on individuals and expect that our theories will generalize straightforwardly to collectives. However, this doesn’t mean that our theories of individual and collective justification need not be continuous.

Our theorizing about collective justification, just like our theorizing about individual justification, places constraints on our theorizing about justification in general. If a theory of justification is viable for individuals, but fails to carry over to collectives, then it is not a good theory of justification. Nonetheless, historically, the majority of our epistemic theorizing has focused on individuals. And our attributions of epistemic states to individuals are more clearly and consistently meant in a serious and literal way. Thus, it makes sense to start with our best theories of individual justification and identify the theories that carry over most straightforwardly to the collective case.

That will be my strategy here. In the next section I situate normic reliabilism and briefly run through some of its virtues. In the remainder of the paper, I will argue that normic reliabilism carries over especially well to collective epistemology without need for modification.


3. What is Normic Reliabilism?

Normic reliabilism is the view that a belief is justified if and only if, given the manner in which it is held, it would require a deviation from normality for it to be false (see Goldman (1986), Leplin (2007), Smith (2010, 2016), and Graham (2016, 2017) for similar views). This can be thought of analogously to the safety condition on knowledge:

Safety: A belief is safe iff it is true at the nearest worlds at which it is held in the same way.\(^1\)

Our analogous condition is as follows:

Normic Reliabilism: a belief is justified iff it is true at all the most normal worlds at which it is held in the same way.\(^2\)

Two important clarifications are in order. Firstly, there are both statistical and non-statistical notions of normality. Normic reliabilism employs the non-statistical notion. This is the notion seemingly invoked by certain uses of generics such as ‘Tigers have stripes’ or ‘sea turtles live to an old age’.\(^3\) These generics describe what is normal for the kind in question. They are not statistical generalizations. Statistically, most sea turtles die within the first moments of life. Nonetheless, it is normal for sea turtles to live to an old age. I wish to stay neutral with respect to the nature of normality. However, one key feature of normality will play an important role in what follows: As pointed out by Smith (2010, 2016), deviations from normality cry out for explanation in a way that normal situations do not. So, a consequence of normic reliabilism is that a belief is justified only if, given only how it is held, it would cry out for explanation if the belief were false.

Secondly, I’m thinking of ‘the manner in which a belief is held’ as encompassing the aspects of the causal explanation of an agent (or group’s) belief that are attributable to the agent (or group). So, when assessing a belief for justification, we consider the most normal worlds at which the agent attributable aspects of the explanation of the belief remain fixed.\(^4\)

---

\(^1\) See Sosa (1999) and Williamson (2000) for canonical statements of safety.

\(^2\) This formulation is a slight simplification, but it will serve our purposes. In §8.1 I will argue that it must, like safety, be reformulated in terms of relevantly similar beliefs.

\(^3\) See Asher and Pelletier (1997, 2012), and Nickel (2008, 2016) for normality approaches to generics.

\(^4\) Thus, the following natural worry is avoided: Jill has entered a lottery. She believes her lottery ticket will win. Intuitively this belief is unjustified. Unbeknownst to Jill, the lottery has been fixed in her favor. But something goes
4. Why Accept Normic Reliabilism

This paper argues for normic reliabilism on the basis of its virtues as an account of collective justification. However, in the spirit of the methodological commitment to continuity, it is worth briefly pointing out some of the more general considerations in favour of normic reliabilism.

Normic reliabilism is a form of normal worlds reliabilism. As others have pointed out, the main advantage of normal worlds reliabilism over standard process reliabilism is its ability to deal with the new evil demon objection (see Goldman (1986), and Graham (2016, 2017)). Imagine that you have been kidnapped and attached to a machine that causes you to have experiences such as that of sitting at a desk and reading this paper. These experiences are misleading. You are not really reading this paper, you are in a warehouse somewhere hooked up to a machine. So, your visual experiences are unreliable. Nonetheless, you are justified in believing that you are reading this paper. The normic reliabilist is able to explain this as follows: at the most normal worlds at which your beliefs are formed in the same way (i.e. you have an experience as of a computer in front of you) your belief is correct. It is only because of a deviation from normality that you are having this experience in circumstances in which it is not veridical.

Normic reliabilism also dovetails with the thought that an agent is justified only if they have ‘done their bit’ toward gaining knowledge (c.f. Smith (2010 p12)). There are several ways to get from the thought that, if one has done one’s bit toward getting knowledge, then one’s belief will be normically reliable:

1. If one has done one’s bit toward getting knowledge then normally what one has done will be enough for knowledge. Knowledge requires truth. So, if one has done one's bit toward getting knowledge then, given what one has done, one’s belief will normally be true.

2. Knowledge is inconsistent with luck. If one holds one’s belief in a manner that yields error in some normal circumstances, then one’s belief is luckily true. So, if one’s belief is not normically reliable, it will fall short of knowledge. Moreover, since normic reliabilism concerns only the

Although her belief is false, it's falsity cries out for explanation. But the normic reliabilist is not committed to saying that it is justified. At the most normal worlds at which we hold fixed only the agent attributable aspects of her belief formation her belief is false. After all, at the most normal worlds in which she forms her belief in the same way the lottery is not fixed in her favor.
aspects of belief explanation that are attributable to the believer, they will be responsible for their failure to acquire knowledge. So, if a believer has done their bit toward acquiring knowledge their belief will be normically reliable.

3. It has been argued by Peet and Pitcovski (2018) that if a belief is safe in a deviant (i.e. non-normal) way then it will not constitute knowledge. They argue for this on the basis of cases like the following:

TEMP: Temp forms his beliefs about the temperature in the room by consulting a thermometer. His beliefs, so formed, are highly reliable, in that any belief he forms on this basis will always be correct. Moreover, he has no reason for thinking that there is anything amiss with his thermometer. But the thermometer is in fact broken, and is fluctuating randomly within a given range. Unbeknownst to Temp, there is an agent hidden in the room who is in control of the thermostat whose job it is to ensure that every time Temp consults the thermometer the “reading” on the thermometer corresponds to the temperature in the room. C.f. Pritchard 2012 p 260.

Temp’s belief is true, and safe. However, it is safe in a deviant way, and so does not constitute knowledge. Temp’s belief is normically reliable, because if we hold fixed only the agent attributable aspects of his belief formation and place him in normal circumstances his belief will be true. However, suppose we have a belief that is true, and safe, but not normically reliable. Furthermore, suppose that the most normal worlds at which a subject forms their belief in the same way will always be modally close to one another (even if they are not modally close to actuality). Then, in normal circumstances, a subject who’s belief is not normically reliable would not be safe. Thus, if it is safe, this is only because of a deviation from normality. So, it will be safe in a deviant way and will not constitute knowledge. Since a failure of normic reliability is a failure that is attributable to the agent, the failure to acquire knowledge will be attributable to the agent. So, if they have done their part toward gaining knowledge then their belief will be normically reliable.

---

5 Pritchard (2012) provides an alternative diagnosis: The problem with Temp’s belief is that his safe true belief is not the result of his cognitive abilities. Peet and Pitcovski respond that anti-luck conditions should be indexed to explanations of belief, rather than belief forming processes, and that ability approaches are unable to handle this shift. This explanation centric approach is reflected in my formulation of normic reliabilism.
These arguments explain why justification is required for knowledge, even if knowledge is not specifically analyzed in terms of justification.

Another major advantage of normic reliabilism, shared with Smith’s normic approach to propositional justification, is its ability to deal with the lottery paradox (see Smith 2010, 2016 for detailed discussion): Suppose I have entered a fair lottery. The probability that my ticket will lose is very high. Suppose I believe on this basis that I will lose. Is my belief justified? Well, it seems I am equally justified in believing of all the other tickets that they will lose. But if I am justified in believing two propositions then, surely, I am justified in believing their conjunction. So, it seems I am justified in believing that all the tickets will lose. But this can’t be right, since I know that one ticket will win. A natural response is to deny that one is justified in believing that one’s ticket will lose. But this seems to have sceptical consequences: consider the proposition that my son is downstairs. I take myself to be justified in believing this. Suppose the probability that my son is downstairs given my evidence is n. So, a probability of n on one’s evidence must be sufficient for justification. But we can imagine a lottery where the probability of any given ticket losing is higher than the probability on my evidence that my son is downstairs. We can apply the paradoxical reasoning to this lottery, and thus demonstrate that n is not sufficient for justification. So, it would seem, my belief that my son is downstairs is not justified after all. But this can’t be right, I can hear him!

Normic reliabilism resolves the problem by disconnecting justification from probabilistic support: My belief that my lottery ticket will lose is not justified as there are normal worlds at which my lottery ticket wins. After all, were my ticket to win this would not cry out for explanation. However, my belief that my son is downstairs is justified. Given my evidence it would require some deviation from normality for him not to be there.

A similar strategy can be applied to the preface paradox. However, this application brings to light perhaps the most uncomfortable consequence of normic approaches to justification: subjects can be justified in believing arbitrarily improbable propositions. After all, the probability of all of a subject’s beliefs being true is very low, even if each belief is normically supported. So, the conjunction of their beliefs might be be normically supported, and thus justified, despite being highly improbable. See Smith (2016 (esp ch 4)) for an extended defense of this consequence.

Dutant and Littlejohn (2020) argue that the preface paradox is problematic for normic conceptions of justification in another way: They ask us to consider the following case:

Eye Exam. An eye doctor has you look into a device that flashes slides. You were asked to identify the number or letter that you see. In each case, the letter or number appears quite clearly to you. The exercise is a bit tedious as it goes on for quite some time without much by way of variability. The letters and numbers aren’t particularly small or blurry. Your doctor says that you did very well and won’t need glasses… Suppose, however, you press your
These do not exhaust the reasons for accepting normic reliabilism. However, they are sufficient to establish that normic reliabilism has a great deal going for it as an account of justification in general. My aim for the remainder of the paper is to provide further support by demonstrating its virtues as an account of collective justification. That is, I will be arguing for the following:

Group Normic Reliabilism: a group belief is justified iff it is true at all the most normal worlds at which it is formed in the same way.

I will consider the advantages of Group Normic Reliabilism (henceforth ‘GNR’) from (what I take to be) least compelling to most compelling. I will begin by arguing that GNR, unlike many other approaches to collective justification, is neutral with respect to the nature of collective belief. For those of us who are not antecedently invested in any particular view of group belief, this is a virtue. Next I will explain how GNR resolves the problem of highly reliable groups with highly unreliable members. Finally, I will turn to the problem of group belief manipulation, arguing that a number of approaches that purport to avoid the problem (Lackey (2016), and Silva (2019)) fail to do so, and that GNR succeeds.

5. Neutrality

They suggest that this is a case in which one is justified in believing an inconsistent set of propositions: You are justified in retaining each of your individual judgments about the numbers you were shown. But you are also justified in believing that the conjunction of these beliefs is false. Normic conceptions of justification entail that you cannot have such an inconsistent set of justified beliefs: your individual judgments regarding the letters and numbers are no longer justified. Initially this seems counter intuitive. However, as Smith (2020) points out, this is actually quite a natural judgment:

“If I have equal justification for each of a series of claims and I learn that exactly one of them is false, then the situation that I confront is very similar to that presented by the lottery paradox. Both situations involve a large set of propositions, each of which is very likely to be true but one of which is sure to be false. In both situations, the propositions are on a par, in that any one could be the false proposition just as easily as any other. Indeed, if we wrote each of the claims down on slips of paper, then there would be one ‘winning ticket’, which featured the one false claim, and a multitude of ‘losing tickets’, which featured the true claims. To believe any particular claim would be tantamount to believing, of one particular ticket, that it’s a loser. For a defender of Closure, accustomed to denying that one can justifiably believe, of a single ticket, that it has lost a fair lottery, embracing en masse defeat in this situation is, I think, a small step.” Smith (2020), p 10.
The nature of group belief, like the nature of individual belief, is highly controversial. For example, summativists take group belief to be determined by the beliefs of individual group members via an aggregation procedure (see Quinton (1975), and List and Pettit (2011)). Joint acceptance approaches take group belief to be determined by a joint commitment (see Gilbert (1987)). Interpretationist approaches (such as Tollefsen (2015)) take group belief to be determined by the best rationalizing theory of the group’s behaviour. And functionalist approaches take groups to have propositional attitudes when they are in a state that satisfies the corresponding functional role (see Huebner (2014)). This list is not exhaustive.

If a theory of justification is neutral with respect to the nature of belief then, ceterus paribus, it has a higher probability of truth given our current evidence than a theory of justification that presupposes a particular view of belief.

Most theories of justification come with very few controversial commitments regarding the nature of individual belief. So considerations of neutrality have little relevance in individual epistemology. By contrast, the majority of approaches to collective justification come with some controversial commitments regarding collective belief. Many theorists are quite explicit in their commitments. For example, Dunn (2018 p2-3) explicitly assumes that a group must undertake an actual vote in order to form a belief. Kallestrup (2016) holds that for a group to be capable belief formation certain joint intentions must be in place. And Schmitt (1994) assumes that group belief requires joint acceptance (that is, roughly, a collective act whereby the group members agree to collectively behave as if a proposition is true).

Some approaches lack such explicit commitments, but nonetheless become implausible if certain controversial assumptions about group belief are false. For example, Lackey (2016) doesn’t make any explicit commitments regarding collective belief. Her approach is as follows:

A group G justifiedly believes that p if and only if

(1) A significant percentage of the operative members of G (a) justifiedly believe that p, and (b) are such that adding together the bases of their justified beliefs that p yields a belief set that is coherent.
(2) Full disclosure of the evidence relevant to the proposition that p accompanied by rational deliberation about that evidence among the members of G in accordance with their individual and group epistemic normative requirements, would not result in further evidence that when added to the bases of G’s members’ beliefs that p, yields a total belief set that fails to make sufficiently probable that p.
Lackey, 2016, 380-381.

According to Lackey, it is impossible for a group to justifiably believe p unless a significant percentage of the group’s operative members believe p. This yields bizarre results on non-summativist approaches to collective belief (i.e. approaches according to which a group can believe p without any of its members believing p). For example, Gilbert (1987) holds that a group believes p when they collectively accept that p. Now consider the following example, which is a collective analogue of one of Lackey’s own selfless asserter cases (Lackey (2007)):

MEDICAL ADVICE: A committee of respected doctors has been gathered to give national advice regarding vaccines. They are all aware of the vast amount of evidence for the safety of vaccines. In particular, they all know that the evidence strongly supports the claim that vaccines don’t cause autism. Rather coincidentally, each doctor has a child with autism. Because of this, each member has nagging doubts about the safety of vaccines, and the relationship between vaccines and autism. They each recognize these doubts to be irrational. Yet, they are still each unable to bring themselves to believe that vaccines are safe. Nonetheless, they each keep their doubts to themselves. During their meeting they review all the evidence and on the basis of the evidence they collectively accept that vaccines don’t cause autism.

It is controversial whether the medical committee really believes that vaccines are safe. However, it seems clear that if such collective acceptance is sufficient for group belief, then this belief is justified. Yet Lackey’s account rules this out. After all, it is not the case that a significant percentage of the group members justifiably believe that vaccines are safe. Lackey’s approach is, thus, plausible only on the assumption that group belief requires group member belief.

There are a number of theories of group justification that do retain neutrality. In particular, neutrality is retained by views of the following form: a group is justified when the group believes p, and the group belief satisfies some property F, such that F can be instantiated in ways consistent with a range of
pictures of collective belief. Here is an example: a group belief that \( p \) is justified iff the group’s reasons support \( p \), and the group is epistemically responsible in believing \( p \) on that basis (see Silva (2019) for such an approach). Plausibly the notion of a group’s possessing evidence and forming a belief in a responsible manner can be made sense of on various different conceptions of group belief.

Likewise, GNR is neutral with respect to the nature of collective belief. If joint acceptance accounts of group belief are correct, then a group belief that \( p \) will be justified iff \( p \) is true at all the most normal worlds in which the group collectively accepts that \( p \), and they reach their collective acceptance in the same way. Whereas, if summativist approaches are correct, and we assume a majoritarian aggregation procedure, then a group belief that \( p \) will be justified iff \( p \) is true at all the most normal worlds at which the majority of the group believes that \( p \), and this state of affairs came about in the same way (I will leave it to the reader to spell out GNR for their favourite account of group belief).

So, GNR is continuous with one of our most promising approaches to justification, and unlike many other approaches to collective justification it is neutral with respect to group belief. This is a good start. But it remains to be seen what positive work GNR can do in collective epistemology.

6. Group Justification and Group Member Justification

Reliabilism holds that a belief is justified iff it is formed in a reliable manner. Lackey (2016) points to a problem for reliabilist approaches to group justification: on certain summativist approaches to collective belief, reliabilism allows a group to be justified even when its members are unjustified.

We learn from the Condorcet jury theorem that a group of independent inquirers who are all above 0.5 reliable with respect to \( p \) can reach an arbitrarily high degree of reliability with respect to \( p \) depending on the group’s size. A group whose members are 0.501 reliable could have reliability approaching 1 if it has enough members. As a result, a group belief can be justified, even when the group members are only 0.501 reliable. Call groups that are reliable in this manner ‘Condorcet groups’. The view that Condorcet group beliefs are justified has a number of counterintuitive consequences:

1. A group could gain the status of justifiably believing \( p \) simply by recruiting more independent members who are 0.501 reliable and believe \( p \). Suppose we have a group for whom it would be practically beneficial to justifiably believe \( p \) (regardless of whether or not \( p \) is true). This group
consists of individuals who believe p and are only just above chance reliable. They could, as a
group, come to justifiably believe p by recruiting more members who believe p and are 0.501
reliable. This problem is discussed further in §7.

2. It makes the grounds of the member beliefs irrelevant to the justification of the group belief.
Justification can be achieved independently of the manner in which members ground their
beliefs.

3. Smaller diligent expert groups will end up less justified than large groups of irresponsible
non-experts. Imagine we have a research team that diligently applies sophisticated research
methods to the question of p’s truth, and rationally deliberates before voting. It will inevitably
be less reliable (and thus less justified) than some large group of 0.501 reliable belief formers.7

Insofar as we retain a commitment to continuity, this is a problem for reliabilism in general. It shows
that reliabilist approaches, no matter how plausible they are as theories of individual justification, are
implausible when applied to collectives.

This is where GNR has an advantage over standard reliabilism. When combined with majoritarian
summativism GNR tells us that a group belief that p will be justified iff p is true at all the most normal
worlds at which the majority of the group believes that p, and this state of affairs came about in the
same way. The bases for the individual member beliefs explain why the members believe as they do.
Thus, they are factored into the determination of group justification.

Moreover, GNR predicts that, unlike the diligent expert group, Condorcet groups are not justified.
Given that the diligent expert group rationally bases their belief on strong evidence, it would cry out for
explanation if they were wrong. This is not to say that they could not be wrong. But, at the most normal
worlds at which they form their belief in the same way, they will not be led into error. This is not true of
the Condorcet group. Although this group has a high probability of reaching the truth, it would not cry
out for explanation if they get it wrong.

7 Deliberative groups are not always reliable, even if composed of mostly rational members. Solomon (2006) points out that
there will often be pressure for such groups to reach consensus, and this can lead to group think and information
suppression. We’ll suppose that this is not the case with our research team.
Consider an analogy: We have a coin that is biased such that it has a 0.501 probability of landing heads. The more times we flip the coin the higher the probability that the majority of flips will land heads. With a very high number of flips there is a very high probability that the majority of flips will land heads. Now suppose that we have a large number of such series of coin flips. If we have a large enough number we should expect that, at some point, there will be a series in which the majority of flips land tails. If and when this happens, it will not cry out for explanation.

Here is another way to put the point: Each coin flip is almost as likely to land tails as it is to land heads. It will not be in any way surprising, nor will it cry out for explanation, when a coin flip lands tails. Moreover, each coin flip is independent. Yet, surely the conjunction of two or more independent unsurprising events is not itself surprising or in need of additional explanation. If correct then this will iterate: a long sequence of independent unsurprising events (like a long sequence of heads) will not be surprising, nor will it cry out for explanation (see Smith (2017) for a defence of this reasoning).

We can think of Condorcet’s theorem as treating voters like biased coin flips (only, they are biased toward the truth). As long as they are above 0.5 reliable, the higher the number of voters, the higher the probability that a majority vote will yield the truth. As with the coin flips, it won’t cry out for explanation if and when Condorcet group gets it wrong. Thus, when a Condorcet group forms a belief via majority voting there will usually be some small number of normal worlds at which they form their belief in the same way, and get it wrong. So, according to normic reliabilism, Condorcet groups will not be justified.

---

8 Two natural responses to this reasoning suggest themselves: 1. The surprisingness of an event is a function of its probability. Since a tails is marginally less probable than a heads it is a tiny bit surprising. And the surprisingness of two events is equal to the sum of the surprisingness of each event. If this is correct then a long series of tails will be surprising. This response is not promising. Firstly, there is no intuitive sense in which a tails would be surprising. Secondly, as Smith (2017) convincingly argues, surprisingness is not a function of probability: If it was then any series of fair coin flips would be equally surprising. So, a series of 92 heads would be just as surprising as a series of 46 heads and 46 tails. If the former is surprising, then the latter would be surprising too. 2. It might be suggested that an event is surprising or cries out for explanation if it is a-typical, where typicality is not a feature of the probability of an event, but rather the probability of the properties of that event (see Williams (2008)). A series of 92 heads would be highly atypical, since the property “all heads” has a far lower probability than the property “between 40 and 60% heads”. Sequences with the latter property would be more typical than sequences with the former property. Smith does not consider this response, but he considers something similar, and his reasoning carries over: The only thing that makes the 40-60% heads property more probable is that there are many ways for this property to obtain. The set of outcomes satisfying this property is larger than the set of outcomes that fails to satisfy it. However, there is nothing about these particular sequences that set them apart from a-typical sequences such that their occurrence requires less explanation. We should not be surprised when a sequence from outside of this set occurs (Smith (2017)).
However, not everyone accepts Lackey’s intuition about Condorcet groups. And, if Lackey is wrong, then GNR is in trouble. For example, Dunn (forthcoming) holds that we should simply accept that Condorcet groups can, by virtue of their size (and, thus, their reliability), be more justified than smaller expert groups. I doubt that there is anything I can say to convince such an ardent process reliabilist. Certainly I find the claim that Condorcet groups are more justified than diligent research groups to be highly counterintuitive. However, I think more can be said in favour of treating Condorcet groups as not possessing genuine justification despite their high reliability.

There is an important distinction between testimonial and instrumental knowledge. We form beliefs differently on the basis of testimony, and the output of instruments. Accepting testimony involves taking someone at their word. This involves trust. It involves recognizing the speaker’s free agency, and the responsibility they have undertaken for the truth of their statement. Instruments on the other hand should be treated as mere truth gauges (see Wright (2014)). The distinction between testifiers and instruments brings with it a difference in vocabulary. We typically assess testifiers in terms of epistemic concepts such as justification, knowledge, and trustworthiness. Instruments, on the other hand, we typically describe merely in terms reliability.

This can help us distinguish between cases in which appraisal in terms of knowledge or justification, as opposed to mere reliability, is appropriate. It is appropriate to treat the outputs of deliberative groups as testimony. It is fitting to adopt an attitude of trust (or distrust) toward such groups, and accept their outputs on this basis. If such a group gets it wrong we might hold them responsible, we may expect them to defend their claims etc. (Fricker (2012)). Such an attitude of trust is not appropriate for Condorcet groups. We would not feel betrayed when they get it wrong. It is more appropriate to treat Condorcet groups as truth gauges than it is to take them at their word. Thus, insofar as the vocabulary of knowledge and justification is inappropriate for merely instrumental sources of knowledge, it seems inappropriate for Condorcet groups.

Dunn is not the only one to suggest that individual rationality and collective rationality sometimes come apart. The apparent independence of collective and individual rationality is widely discussed in the philosophy of science (see, for example, Mayo-Wilson, Zollman, and Danks (2011)). The basic point driven home in this literature is that, from the perspective of design, it is better for our epistemic communities to contain members who don’t always behave in an (individually) rational way. Such communities are more prone to producing theories that approximate the truth. My core response carries
over just as well here: large epistemic communities are best assessed instrumentally rather than
agentially. However, we can say more: from the perspective of design, it is rational that our
communities contain members who act irrationally. However, this doesn’t make the communities more
rational. Consider an analogy: because of the limits of human cognition, the sorts of environments we
inhabit, and the types of reasoning problems we typically face, it makes sense that we have certain
biases, and are prone to certain forms of irrationality such as a tendency to commit the base rate fallacy.
From the design perspective, it makes sense that we have these biases. It makes us more reliable
overall. However, beliefs formed via the base rate fallacy are still irrational.

My response to the condorcet problem could also be attacked from another angle: Imagine that we
discovered that one of our core cognitive faculties works as follows: there are many nodes that respond
to inputs by selecting between a number of representational outputs. These selections are akin to votes,
and the system as a whole outputs the representation with the most votes. Individually these nodes are
not very reliable (perhaps just above chance), but that there are so many of them that, together, they are
highly reliable. Would we not want to say that beliefs based on the outputs of this system are justified?

Such a discovery would certainly be problematic for GNR. However, it would also spell trouble for far
less controversial epistemic views, such as the view that safe belief is required for knowledge. After all,
such lottery like systems would not only yield falsity at some normal worlds, but also at some nearby
worlds. Similarly, we would be forced to revise the assumption that we cannot know that our lottery
ticket will lose merely on the basis of the high probability that it will lose. And we would have to revise
the assumption that knowledge requires an explanatory connection between our beliefs and the facts.
So, GNR is in good company here. Until we have reason to suspect that our cognitive systems actually
do work this way, we do not have strong reasons to reject GNR.

Up to this point we have seen that GNR is continuous with one of our most promising theories of
justification. It is neutral with respect to the nature of group belief. And, unlike other reliabilist
approaches, it is able to make the correct predictions about Condorcet groups. Still, many theories of
group justification are continuous with promising theories of individual justification. Many theorists of
group justification are invested in particular approaches to collective belief, and thus will not view
neutrality as a virtue. And the Condorcet problem only arises for reliabilism. In the next section I will
discuss a more general problem for theories of collective justification. I will argue that even those
views designed to resolve the problem fail. But, where they fail, GNR succeeds.
7. Group Belief Manipulation

Beliefs formed or sustained either at will or through intentional manipulation of one’s own belief system for purely practical purposes (i.e. not for epistemic reasons) are not justified. Any approach to justification that counts as justified a belief formed or sustained in this manner will be untenable. As Lackey (2016) argues, many approaches to group justification allow groups to gain (or sustain) justified beliefs in this manner.

The problem applies most obviously to joint acceptance accounts of collective justification. These accounts hold that group justification depends on the reasons possessed by a group, and that a group to possess a reason only when it is mutually accepted by the group members (see Schmitt (1994)). The problem is, if a group wants to believe p, and there is evidence against p, they can simply refuse to collectively accept that evidence. Through such manipulation of evidence groups can form justified beliefs in almost any proposition they choose. For example, Philip Morris can justifiably believe that smoking is safe by failing to accept evidence for the dangers of smoking (Lackey (2016)).

The manipulation problem can be generalized to a range of other views. For example, on most summativist views a group can come to justifiably believe p by inviting new members who justifiably believe p. Consider the following approach: a group justifiably believes p iff the majority of its members justifiably believe p. Now imagine that the majority of Philip Morris members do not justifiably believe that smoking is safe. They can bring it about that they (i.e. Philip Morris) justifiably believe that smoking is safe by inviting new group members who possess no evidence regarding the dangers of smoking, and telling them upon joining that smoking is safe.9

Lackey intends her own approach to avoid the manipulation problem. Her account is as follows:

GEAA: A group G justifiably believes that p if and only if

9Dunn (forthcoming) responds on two grounds. Both responses interpret Lackey as holding her opponents to the claim that Philip Morris isn’t justified in believing e.g. that smoking is dangerous. He interprets Lackey as assuming that Philip Morris does justifiably believe that smoking is dangerous. His first response is that reliabilists can claim that Philip Morris had propositional but not doxastic justification for believing that smoking is dangerous. The second response is that we only want to be able to say that Philip Morris is justified in believing that smoking is dangerous in order to hold them responsible for their actions. But, as he points out, we can hold them responsible without attributing such a justified belief. These responses are unconvincing, as they fail to target what is really at issue: the fact that groups can generate doxastically justified beliefs in whatever proposition they choose. Neither of Dunn’s responses touch on this.
(1) A significant percentage of the operative members of G (a) justifiedly believe that p, and (b) are such that adding together the bases of their justified beliefs that p yields a belief set that is coherent.

(2) Full disclosure of the evidence relevant to the proposition that p accompanied by rational deliberation about that evidence among the members of G in accordance with their individual and group epistemic normative requirements, would not result in further evidence that when added to the bases of G’s members’ beliefs that p, yields a total belief set that fails to make sufficiently probable that p.

Lackey, 2016, 380-381.

GEAA seemingly avoids the manipulation problem because in order to gerrymander the group membership in such a way as to guarantee justified belief in p, the person doing the gerrymandering must be aware of defeaters for p. But, if they are aware of the defeaters, then the group will fail condition (2).

Lackey acknowledges that group membership could be manipulated from outside. But she argues that this is unproblematic: it is akin to a normal agent being fed misleading evidence from outside. If an external party misleads us then we will form justified false beliefs. No epistemic wrongdoing will be attributable to the believer in these circumstances. The wrongdoing is attributable to the one who manipulates the evidence. Likewise, when group membership is manipulated from outside, the epistemic harm is not attributable to the group.

Unfortunately, GEAA still faces a version of the manipulation problem:

GENIE: Suppose Sally and Ben run a company that produces Xs. They falsely believe that X’s are safe. They each believe this for a different reason, and each reason is defeated. So each is unjustified. However, they each possesses a defeater defeater for the others’ defeater. So, if combined their total evidence would actually provide some support for p. Moreover, it suits them to believe that X’s are safe. And they desire for practical reasons that the company justifiably believes this.
They decide to hire some directors and managers for the company. In the night they are visited by a genie. The genie tells them the following: ‘I am granting you each new powers of discrimination. From now on you will, with a certain form of reliability (to be specified), be able to tell whether somebody else justifiably believes that X’s are safe, in such a way that the basis for their belief is consistent with the basis for your own beliefs (and those of your employees), and they possess no defeaters for the basis of your beliefs (or those of your employees). There will be no false positives (i.e. you will never falsely believe somebody satisfies these conditions), but there will be a random rate of failure to detect agents that do satisfy these conditions. You will not be aware of this rate of failure’.

Using their new powers they start interviewing candidates. They only hire people who justifiably believe that X’s are safe. They do so simply because it will be beneficial for the company if everyone justifiably believes X’s are safe (regardless of whether or not they actually are safe).

They face many candidates who don’t appear to justifiably believe that X’s are safe. But they have no idea whether this is due to the low reliability of their new faculty, or rather whether it is due to the low number of people who actually justifiably believe X’s are safe.

Consider the resultant group: almost all of the operative members justifiably believe that X’s are safe, and their bases are consistent. Moreover, full disclosure of the group’s evidence would not throw up any defeaters with respect to the safety of X’s (after all, Sally and Ben both possess defeater defeaters for the others’ defeaters). Thus, according to GEAA the group justifiably believes that X’s are safe. This is the wrong result. The group membership was intentionally and illegitimately manipulated from within in order to ensure that the conditions for GEAA were satisfied. This manipulation of membership is attributable to the company. Thus, this is not a case of justified group belief.

GENIE also spells trouble for the only other account of collective belief I am aware of that is designed to avoid the manipulation problem: Silva’s (2019) evidentialist responsibilism:

A group, G, justifiably believes that P on the basis of evidence E iff:
(1) E is a sufficient reason to believe P, and the total evidence possessed by enough of the operative members of G does not include further evidence, E*, such that E and E* together are not a sufficient reason to believe P, and

(2) G is epistemically responsible in believing P on the basis of E.

Silva’s first condition is clearly satisfied in GENIE. It may seem that the second condition isn’t satisfied. After all, it certainly doesn’t seem that the group is being responsible. However, Silva spells out responsible group belief in terms of ‘enough’ of the individual members of G responsibly fulfilling their epistemic duties qua members of G. In GENIE, almost everyone in the group justifiably (and so responsibly) believes X’s are harmless.

Perhaps Sally and Ben are not living up to their epistemic duties. After all, their beliefs are not justified, and they have manipulated the group’s membership. However, Sally and Ben constitute a very small subset of the total group. So, it is still plausible that ‘enough’ of the members have been responsible (unless ‘enough’ needs to be 100% of the membership). Despite the majority of group members being responsible, the real reason that the group as a whole (as opposed to its members) believes p is that this is pragmatically beneficial. Given the way the group was constructed it would inevitably end up believing p, even if the reasons possessed by the individual group members were different. So the group itself is not justified. There may be other ways of thinking about group epistemic responsibility that avoid this problem. However, it is not clear how this would be done whilst maintaining neutral on the nature of group belief.

So, manipulation is a problem for a great many approaches to collective justification.\(^\text{10}\) However, it is not a problem for GNR. The reason for this is simple: In cases of evidence/membership manipulation the process of specifically selecting evidence or group members so as to reach a particular verdict for purely pragmatic reasons is a large part of the explanation for the group’s continued belief. The beliefs are formed and maintained for practical, not epistemic reasons. As Hughes (MS) puts it whist discussing a related issue, beliefs so formed will be ‘indifferent and insensitive to their environments’. Thus, it will not cry out for explanation if and when such beliefs are found to be false.

\(^{10}\) The only approach I am aware of that avoids the issue is Kallestrup’s (2016) group virtue epistemology. If one is is antecedently sympathetic to virtue reliabilism (I am not, and have argued against the virtue reliabilist framework in general elsewhere (XXXX)), and one is already invested in non-summativism about of group belief (and, hence, does not feel the pull of neutrality), then group virtue reliabilism also constitutes a promising approach to collective justification.
Consider joint acceptance accounts of collective belief: Suppose a group collectively accepts p. If the resultant belief is to be justified then, given the process leading up to this act of collective acceptance, it must cry out for explanation if p is false. If this process involves the intentional restriction of evidence, or the ignoring of evidence among group members, then it will be no surprise if the group’s belief ends up being false. Similar reasoning applies to other approaches to group belief.

Thus GNR, unlike almost every other extant view of group justification, avoids the manipulation problem. It is, thus, a promising approach to group justification. It is theoretically elegant, continuous with our theorizing about individual justification, neutral as to the nature of group belief, and avoids the key problems facing its closest rivals. Given the importance of continuity between the individual and collective realm, this counts strongly in favour of normic reliabilism as a view of justification in general. In the final section I outline and respond to two objections to normic reliabilism.

8. Objections and Replies

8.1 Irrational yet normically reliable

The first objection derives from Gardiner’s (2016) discussion of statistical evidence in law. Although not presented directly as an objection to normic reliabilism, it is problematic for normic reliabilism as currently formulated. Gardiner asks us to consider the following case:

Oliver believes p, where p is “Jones murdered Jill.” He believes this because Jones’s clothes are bloodstained. Oliver’s belief is thus based upon evidence that normically supports p. But Jones believes p because he is confident that only butchers wear bloody clothes and butchers are angry and murderous.

Gardiner, forthcoming, p 15.

Suppose that at all the most normal worlds at which Jones has a bloody shirt he is the murderer. If this is the case then, since Oliver believes that Jones is the murderer partly on the basis of his bloody clothes, his belief is true at all the most normal worlds at which it is formed in the same way. So, normic reliabilism must say his belief is justified. Yet, his belief is clearly not justified. So, normic reliabilism must be rejected.

This objection is fatal to normic reliabilism as currently formulated. However, as I noted earlier, the formulation we have been operating with is a simplification. Normic reliabilism is modelled on the
safety condition on knowledge, with normal worlds replacing nearby worlds. The corresponding version of the safety principle is as follows:

Safety: A belief is safe iff it is true at all the most nearby worlds at which it is formed in the same way.

This version of the safety principle is untenable. For example, it can’t deal with luckily true beliefs in necessary truths. The standard response is to hold that, in order for a belief to be safe, the basis upon which it’s formed must not lead the believer to error in any nearby worlds. That is, there cannot be any nearby worlds in which the believer forms a (relevantly similar) false belief on the same (or, a relevantly similar) basis (see Williamson (2000), Pritchard (2009), Manley (2012), and Grundmann (forthcoming)).

If we wish to retain parity between safety and normic reliability normic reliabilism must be reformulated as follows:

Normic Reliabilism: a belief is justified iff, at the most normal worlds in which the agent forms a similar belief in a relevantly similar manner (e.g. on the same basis), the agent is not led to error.

This new formulation of normic reliabilism avoids Gardiner’s worry. After all, there are many other beliefs that Oliver could easily form on the same basis that will be false in some normal worlds. For example, given that Oliver infers from Jones’s bloody shirt that he is a butcher, it would be natural for him to also infer that he is not a vegan, and that he owns some sharp knives. These beliefs are liable to be false at some of the most normal worlds at which Oliver forms his belief in the same way. Thus, Oliver’s belief that Jones is the murderer is not normically reliable.

8.2 Easy Defeat

Backes (2019) has objected to the family of views to which Normic Reliabilism belongs. Backes argues that these views suffer a problem of easy defeat: our otherwise justified beliefs can be defeated by highly improbable possibilities. Consider the following example:

PEANUTS: Helen is allergic to peanuts. She goes to a cafe’ and orders a brownie labeled ‘peanut free’. Based on this Helen is justified in believing P, that the brownie is safe to eat. On
the Normalcy View this entails that there does not exist a single normal world in which Q, the brownie has been contaminated by something containing peanuts. After ordering the brownie Helen sees a newspaper headline that reads D, international flour supplier admits to having accidentally put 1 bag of peanut-contaminated baking flour into circulation.

Backes 2019, 2884.

Intuitively, Backes holds, Helen is still justified in believing that her brownie is safe to eat. However, normic reliabilism seems to entail that her belief is defeated: we might suppose that Helen forms her belief by simply following her evidence (and that she does so in a responsible manner). If this is the case, then her belief will be justified iff it is true at the most normal worlds consistent with her evidence (just as Smith’s (2016) normic account of evidential support predicts). However, once she has acquired additional evidence (i.e. after she has read the newspaper headline) her belief is no longer true at all the most normal worlds consistent with her evidence. So, it might be thought, her belief is no longer justified.

There are two key points to note in response. Firstly, it is not clear that the normic reliabilist is committed to holding that Helen’s belief is defeated. Recall that Normic Reliabilism relativizes the relevant worlds to those at which the agent holds their belief in the same (or a relevantly similar) way. That is, we consider the truth of the belief at the most normal worlds at which the agent attributable elements of the explanation for the agent’s belief remain the same. This raises the question: does Helen’s reading of the newspaper headline alter the causal explanation of her belief? If, in the face of her new evidence, she continues believing that the brownie is safe, does this in any sense require additional explanation?

Suppose the answer is no: her ignoring this new information does not enter into the causal explanation of her continued belief (this will be true if normal well functioning epistemic agents would not normally alter their beliefs in response to such evidence), then Normic Reliibilism predicts that Sally’s belief continues to be justified. After all, if her ignoring the evidence does not change the explanation for her continued belief, then she continues to believe in the same manner as before. And, at the most normal worlds at which she continues to believe in the same manner, her belief is true. This is in-spite of the fact that, upon reading the newspaper headline, she realizes that the actual world is not amongst these most normal worlds.
That said, I doubt that a normal well functioning epistemic agent would continue believing in Helen’s position. I suspect that her continued belief does cry out for explanation. Put yourself in Helen’s shoes: you are staring down at your brownie wondering ‘will this kill me?’. Before reading the headline you would likely think ‘no, it won’t kill me, there are strict practices in place to prevent food contamination’, but after reading the headline it would be quite natural (and, I believe, epistemically proper) to weaken one’s position, thinking ‘it might kill me, but it is extremely unlikely to do so’. However, if this is right, Helen’s belief is defeated. She should not maintain an all-out belief that the brownie is safe. Rather, she ought to merely believe that it is highly probable that the brownie is safe. This belief is still justified. After all, in all the nearest worlds in which her belief is formed and sustained in the same way it is highly probable that her brownie is safe.\(^{11}\) In such cases we may still describe Helen as believing that her brownie is safe. After all, it is very unlikely that her behaviour will change in any major way upon revising her belief. But strictly speaking this would be false. If she is epistemically responsible, she will jettison her all-out belief (Smith (forthcoming) responds similarly).\(^{12}\)

**9. Conclusion**

I have argued that our theories of justification should be continuous between the collective and individual realm. We can test our best theories of individual justification to see how well they carry over to the collective realm. If a theory of justification carries over straightforwardly to the collective realm, this is a point in its favour. If it is able to do substantial explanatory work in collective epistemology, even better. I have argued that normic reliabilism does particularly well on this score: GNR, which is fully continuous with normic reliabilism, is currently our most promising view of collective justification. I argued for GNR on two main grounds: Firstly, unlike most approaches to collective justification, it is simultaneously continuous with our best theorizing about individual justification, and neutral with respect to the nature of group belief. Secondly, it is able to solve a number of recalcitrant problems, some of which are faced by all of its rivals. This provides strong support for Normic Reliabilism not only as a view of collective justification, but as a view of justification in general. I have also defended Normic Reliabilism from recent challenges by Gardiner (forthcoming) and Backes (2019).

---

\(^{11}\) See Moss (2018) for an account of the conditions under which such a probabilistic belief is true at a world.

\(^{12}\) See also Nelkin (2000) who provides a similar response to a similar worry raised by Harman (1986).
Bibliography

Author (XXX)
Hughes, N. Manuscript. Epistemic Feedback Loops.


