View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Edinburgh Research Explorer

THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Expressivism and collectives

Citation for published version:
Ridge, M 2018, 'Expressivism and collectives', *Mind*, vol. 127, no. 507, fzw078, pp. 833-861. https://doi.org/10.1093/mind/fzw078

Digital Object Identifier (DOI):
10.1093/mind/fzw078

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Mind

Publisher Rights Statement:
This is a pre-copyedited, author-produced PDF of an article accepted for publication in Mind following peer review. The version of record Michael Ridge; Expressivism and Collectives, Mind, , fzw078, https://doi.org/10.1093/mind/fzw078 is available online at: https://academic.oup.com/mind/article/4036429/Expressivism-and-Collectives.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 22. Feb. 2020

Expressivism and Collectives[1]

Abstract:

Expressivists have a problem with collectives. I initially illustrate the problem against the background of Allan Gibbard's expressivist theory, where it is especially stark. I then argue that the problem generalizes. Gibbard's account entails that judgments about what collective agents ought to do are contingency plans for what to do if one is in the circumstances facing the relevant collective agent. So, for example, my judgment that the US ought not to have invaded Iraq is a contingency plan for what to do if I find myself in the circumstances facing the US on the eve of the Iraq war. I argue that, given plausible functionalist assumptions in the philosophy of mind expressivists would struggle to reject, such contingency plans are *impossible* for individual agents. Since normative judgments about collectives are possible, Gibbard's theory is unacceptable. Having demonstrated the basic problem for Gibbard's view, I then argue that the problem has a much wider scope. In particular, I argue that traditional forms of expressivism cannot simultaneously provide a plausible account of *both* agent-relative normative judgments *and* judgments about what collectives ought to do. Fortunately, all is not lost; hybrid or "Ecumenical" form of Expressivism can disarm the dilemma, or so I argue. I conclude by briefly sketching some further challenges Ecumenical Expressivism must overcome for its treatment of judgments about collectives to be fully satisfactory and indicate possible lines of argument to be pursued in future work on these challenges.

[1] Thanks for helpful comments to Matthew Chrisman, Guy Fletcher, Allan Gibbard, Sebastian Kohler, Elinor Mason, Debbie Roberts, two anonymous referees for *Mind* and audiences at the Global Expressivism Conference in Szczecin and the New Directions For Expressivism Conference in Sheffield. I am especially grateful to Matthew Chrisman as it was a conversation with him which gave me the idea for this paper in the first place.

Expressivism and Collectives

Expressivists aspire to make sense of ordinary normative thought and discourse. They therefore need an account of normative judgments about collective agents, as such judgments are commonplace. People routinely express their moral outrage at the actions of nation-states, corporations and non-governmental organizations. Collective agents are often held responsible for their actions. On the eve of the US invasion of Iraq, people said things like "The US ought not to invade Iraq." Purely prudential normative judgments about collectives are also frequently made. Anti-war activists and foreign policy experts said things like, "The US should not invade Iraq – overextending our military would be a strategic blunder." Ordinary people argue about what corporations like Microsoft ought to do about some problem with their software, what direction a Philosophy department ought to take in its next set of hires, whether the Mets should trade one of their players, and so on.

Even though such judgments are commonplace, there has so far been no discussion of what an expressivist should say about them. One might be forgiven for assuming this is because expressivism faces no distinctive problems from this quarter, but this is not the case. My initial presentation of the problem arising here focuses on Allan Gibbard's theory in *Thinking How to Live*. With this problem on the table, I argue that expressivists more generally struggle to deal with such judgments while at the same time adequately making sense of agent-relative normative judgments (section I). I then consider what I take to be the best replies to this problem within the framework of a traditional form of expressivism and argue that they fail (section II). Finally, I argue that a hybrid or "Ecumenical" form of

Expressivism can disarm this dilemma, providing a satisfying account of normative judgments about collectives while still making good sense of agent-relative normative judgments (section III). I conclude by discussing residual problems for this solution.

I. The Problem

I first clarify what I mean by 'expressivism', as this term has been used to refer to a variety of views. By 'expressivism' I mean a view which explains the meanings of items in a target area of discourse by adverting to distinctive states of mind expressed by claims made with that discourse. Moreover, these states of mind are not to be understood, in the first instance, in terms of representing the world as being a certain way. In the context of normative discourse, the states of mind are instead traditionally understood in the first instance as desire-like, or anyway at least partially constituted by desire-like states. On some versions of expressivism, it is a view entirely within "meta-semantics," but on other versions of the view it also takes on non-trivial commitments in semantics.[2] For purposes of this paper, I can remain neutral on this point. I am, however, excluding views which locate their expressivism entirely in pragmatics.[3] Expressivism in my sense therefore must provide an account of what is distinctive about normative judgments, and then explain the meanings of normative sentences in terms of expressing such judgments.[4]

---

[2] For purely meta-semantic interpretations of expressivism, see e.g. Charlow 2013, Chrisman 2013, Chrisman 2016, Ridge 2014, Sepielli 2012, Silk 2013, and Sinclair 2016. For an argument that expressivism must take on commitments in semantics see Schroeder 2008 and Schroeder 2010. See also Alwood 2015. Yet another approach is to understand expressivism in terms of the broader framework of dynamic semantics. See, e.g. Charlow 2015. Compare also Perez Carballo and Santorio 2016.

[3] Seth Yalcin, for example, understands expressivism as first and foremost a pragmatic thesis, albeit one which imposes constraints on the semantics for the relevant fragment of discourse. See, e.g. Yalcin 2012. Yet another view that goes under the heading of 'expressivism' is that the target words do not contribute to semantic content at all, but instead act as speech-act modifiers. For a view of this kind with regards to epistemic modals, see Schneider 2010. This sort of speech-act modifier view (or something close to it) can also be seen in many of the classical ethical expressivists, perhaps most notably in Hare's work; see Hare 1952. See also Stevenson's account of 'good' and in particular his 'do so as well!' clause in e.g. Stevenson 1944. Compare Ridge 2003 and Ayer 1952, chapter six. More recently, some "hybrid theorists" who call themselves expressivists locate their expressivism in pragmatic phenomena like the implicature of a non-cognitive attitude, though in my view it is confusing to refer to such views as "expressivist" since in an important sense they are more aligned with traditional forms of cognitivism. See, e.g. Barker 2000.

[4] In this respect, expressivism in my view stands in contrast not only with some of the pragmatic accounts listed in the previous footnote, but also with inferentialist theories of meaning which are also often referred to as forms

I initially focus on Allan Gibbard's theory in *Thinking How to Live* (Gibbard 2003). Gibbard there provides an account of judgments about the "thing to do" – more commonly glossed as the all things considered 'ought'. Since Gibbard's theory provides one of the most impressive and systematically worked out versions of expressivism in the literature, demonstrating that the problem is a stark one for his view should already be of some interest. Gibbard reminds us that ethics "concerns what to do," and argues that various forms of moral realism "drive what to do out of ethics."[5] Suppose reductive naturalism were true. Whatever natural property N the naturalist privileges as normative, one can intelligibly admit that an action is N but still wonder whether it is the thing to do. Nor does non-naturalist realism obviously do any better. Clearly this is not the place to rehearse these familiar arguments and their many epicycles.[6] Gibbard argues that judging that an action is the thing to do in a given circumstance just is *planning* to perform it if in that circumstance, and that this explains how normative judgments settle the thing to do.

Of course, we make normative judgments about what *other* people ought to do, but Gibbard has a ready account of this. In addition to forming plans for circumstances we anticipate facing, we form contingency plans for circumstances we "fully know we may never face." My judgment about what someone else should do is constituted by my planning what to do if in circumstances *precisely like* those facing that person. Here is Gibbard:

> What is Holmes doing when he thinks that in Mrs. Hudson's situation, the thing to do is to project the silhouette and stay in?...Holmes is planning, hypothetically, we might

---

'of expressivism'. In this case, the idea is that we understand meaning in terms of the expression of an inferential role, rather than in terms of the expression of a state of mind. See Brandom 1994. For a discussion of how Brandom's inferentialist form of expressivism might fruitfully be combined with expressivism in Blackburn's sense, see Price 2011.

[5] Ibid., pp. 13-20.

[6] Because Gibbard's argument puts so much weight on what is "intelligible," it might seem not to apply to so-called "Cornell Style" realism, but there are familiar ways of extending this general strategy to target those views as well. Readers well stepped in these debates already will be relieved to know that we will not be visiting Cornell, Moral Twin Earth or the island on which Hare's missionary met the cannibals, though all of those vistas would be essential first steps for anything like a proper survey of the relevant terrain. Those not familiar with this territory might begin with Boyd 1988, Horgan and Timmons 1992, and Hare 1952.

> say, what to do if in Mrs. Hudson's shoes…How much of her plight is Holmes planning for, hypothetically, when he comes to a view as to what the thing for her to do is? *All of it, I answer.* (Gibbard 2003: 50)

Obviously more could be said about Gibbard's view, but this is enough to set up the basic problem. Consider a seemingly obvious corollary of Gibbard's view when it comes to judgments about collective agents. I judge that US ought not to have invaded Iraq in 2003. On Gibbard's account, this judgment is constituted by a contingency plan for what to do if in circumstances *precisely like those facing the US on the eve of the 2003 invasion of Iraq.* These include being a massive nation-state that can declare and wage war. This is just plain weird. Granted, I may plan for all sorts of wild possibilities. I may, in the right frame of mind, deliberate about what I would do if I won the lottery or even if I were Superman. However, forming a contingency plan for what to do if I were in precisely the circumstances facing the US on the eve of the Gulf war is radically different. Explaining precisely *how* the case of collectives differs relevantly from these other cases is not trivial, though.

It is important at the outset to put an initially tempting objection to one side. One might appeal to the intuitive weirdness of this corollary of Gibbard's theory to argue that it contradicts the phenomenology of normative judgment and should therefore be rejected. This objection is unconvincing. Expressivists must anyway argue that the phenomenology of normative judgment, insofar as there is such a thing, can be misleading. After all, it probably does not seem "from the inside" that I am planning what to do if in circumstances just like those facing Superman when he is facing down a giant robot when I judge he ought to use his heat vision. Indeed, it is part of the *point* of expressivism that normative judgments are different from factual judgments in ways that are superficially invisible. Rather than drawing plausibility from its fit with the phenomenology of making normative judgments, expressivism draws plausibility from its explanatory power. Objecting to expressivism on

these grounds is no more promising than objecting to functionalism on the grounds that it does not seem "from the inside" that my belief that it is raining outside is such-and-such functional state.[7]

A better objection begins with a weaker transparency assumption. Granted, we have no robustly privileged access to the underlying nature of our judgments. However, the following much weaker transparency thesis about intention is plausible, as Richard Holton has argued in Holton 1993:

> For all rational agents X, unless we have contrary evidence we are entitled to believe: X intends to Φ if and only if X believes that he intends to Φ.

I here briefly rehearse Holton's argument. According to functionalism, we characterize a given mental state with a set of biconditionals which specify a functional role associated with that state. Someone who is in the state need not satisfy *all* of these biconditionals, but he must meet enough of them or we lose our grip on the idea that he occupies that state. How much is enough? There is no precise answer to this question, which raises broader questions about functionalism I cannot explore here. What *are* the biconditionals which characterize intending to Φ? Holton plausibly suggests that one such biconditional is that people intend to Φ if and only if they believe they intend to Φ.

Obviously there is a lot more that could be said about this, but we are already in a position to see how this might frame an objection to Gibbard's theory. Because Gibbard's theory *identifies* judgments about what one ought to do with intentions (I assume that 'intention' and 'plan' are interchangeable), it would be a problem for him if people who judged that someone ought to Φ in C did not also in general believe that they intended to Φ if

[7] For discussion of transparency assumptions and expressivism more generally, see Woods 2014, Toppinen 2014 and Ridge 2009. See also Horgan and Timmons 2006 and Joyce 2010. I discuss the analogy with functionalism more below.

they found themselves in circumstances C.[8] Note how this differs from the problematic argument from phenomenology I sketched above. The person making the judgment does not need to realize that *in judging that A ought to* Φ in C he *thereby* formed an intention to Φ in C. He need not realize Gibbard's theory is true, after all. Rather, having successfully formed what is *in fact* an intention to Φ in C, he must believe he intends to Φ in C. Or at least, there is a presumption that he so believes; otherwise he must satisfy enough of the other biconditionals.

This is also plausible as applied to the case of our normative judgments about other individuals, including individuals in radically different circumstances from our own. Suppose I assert that Frankenstein ought not to have created his monster. You are not sure you heard me correctly, and so you ask, "Are you saying if you were in Frankenstein's shoes, you wouldn't have created the monster?" It certainly would not be weird or surprising for me to say something like, "That's right – although I don't expect it ever to come up!"

Unfortunately for Gibbard, the corresponding view *does* sound crazy when transposed to judgments about collectives:

Jones: "The US ought to invade Iraq."

Smith: "Are you saying that you plan, if in circumstances just like those the US is in right now, to invade Iraq?"

Jones: "What? No, of course not. What a weird thing to ask! I don't have any plans for what to do if I were a nation-state! Obviously it is impossible for me to be in those circumstances. You must think I have bizarre delusions of grandeur!"

---

[8] In *Thinking How to Live* Gibbard presents himself, at least initially, as merely offering a "possibility proof" that normative language could work as his expressivist theory predicts – that such a language is at least possible. It is clear, though, that Gibbard is ultimately trying to convince the reader that this is not only possible, but that in fact it is the best construal of our actual discourse. Insofar as this is correct, then we should, at the end of the day, *identify* normative judgments with the relevant planning states. This is fairly clear from the *Preface* where he remarks, "The hypothesis of this book is easy to state: *Thinking what I ought to do is thinking what to do."* (x) Given that thinking what to do just is planning what to do, it is clear that Gibbard's ultimate conclusion is about the nature of normative judgment itself. This conclusion is therefore presumably modally robust – that is, what Gibbard really wants to establish is the *essential* nature of normative judgment.

In the case of judgments about other individuals in weird circumstances, it is not crazy to think people who make such judgments in a normal sort of case will indeed believe that they intend to act as they judge the individual in question ought to act *if* they were ever in identical circumstances. By contrast, in the case of our judgments about collectives, we do not expect people making such judgments to believe they plan to act in a certain way if they turn out to be in that collective's exact circumstances. Indeed, we rightly think that normal people do not form intentions about what to do if they were collectives. One important piece of evidence for this, given functionalism, is that they do not believe they have any such intentions. In particular, they do not believe this in circumstances which Gibbard's theory plus the functionalist theory predicts they would.[9]

Of course, someone might intend to do what they can to bring it about that such a collective would act as they judge it ought. If I judge the US invasion of Iraq would be immoral then I might well intend to do what (little) I can to prevent this from happening. This, though is simply not the same as judging that the US ought not to invade. As I explain below, given that our normative judgments are sometimes agent-relative, one can coherently judge both that the US ought not to invade Iraq *and* at the same time judge that the President ought to bring it about that the US invades Iraq. Given a theory like egoism, e.g., the ends the President ought to pursue (popularity) and the ends the US ought to pursue (the national interest) can come apart.

One might worry that because contingency plans by individuals for what to do if they were collectives are so weird, the normal connection to believing one so intends is

[9] One might object that in these cases we have immediate evidence that misleads us about our own intentions, namely the mere fact that the relevant circumstances involve being a collective agent, which is weird even if possible. If it were just the weirdness of the intention, though, then we would be just as quick to balk at the attribution of contingency plans for what to do if weird individual agents, which we don't – or anyway, not to the same extent. Moreover, the fact that the circumstances include being a collective agent presumably would reliably count as good evidence that the individual has no such contingency plan only if it is impossible to form intentions for such contingency. In that case, though, Gibbard's theory is sunk anyway. Thanks to an anonymous referee for pressing me on this point.

systematically broken. This reply is somewhat ad hoc, without some deeper explanation of the supposed systematic failure of this particular biconditional. In fact, the challenge can be sharpened in a way that further undermines this reply.

Consider what the *other* biconditionals might partially define the functional role of intention. Here is a very plausible one:

> For all rational agents X, X intends to Φ in C if and only if X would try to Φ if he believed himself to be in C.

It is hard to imagine a plausible functionalist treatment of intention which did not somehow advert to the role of intention in controlling action. Of course, it is not constitutive of intention that it leads one to do what one intends. More plausibly, intention will normally lead an agent at least to *try* to act accordingly when the opportunity presents itself. We do perhaps sometimes suffer from weakness of will so severe that we do not even try to do what we intend even when the opportunity arises. That, though, is simply to allow that the relevant biconditionals need not be satisfied all the time for a state to count as intention.

The right-hand side of this biconditional should be understood as a counterfactual conditional. If it were construed as a material conditional then it would produce too many false positives - cases in which the antecedent just happens to be false but in which we would not think of the agent as thereby intending accordingly. The counterfactual conditional also fits better with the idea that we can advert to intentions to explain how agents would behave in different merely possible circumstances. If we read the right-hand side in this way, though, we have to ask what the nearest possible worlds in which the agent believed himself to be the relevant collective would be like – and in particular, whether the agent would try to do what he believes the collective ought to do.

Suppose I believe the US should have gone to war with Germany in 1940. On Gibbard's account this means I have a contingency plan to *go to war with Germany* if in

circumstances just like those facing the US in 1940. On the biconditional we are considering, we need to see whether I would try to go to war with Germany in the nearest possible worlds in which I believe I am the US in 1940 – or, more accurately, in which I believe I am in circumstances *precisely* like those facing the US in 1940, including my being a massive nation-state like the US. The problem is that the nearest possible worlds in which I believe this are most likely worlds in which I am suffering from a delusion. Would I try to go to war with Germany in those worlds? It is not even clear what would *constitute* my trying to go to war with Germany, just for a start. Even if we can get around the non-trivial conceptual questions about just what would make it true that some individual is literally trying to go to war with Germany, who knows if I would try to do this in the nearest possible worlds in which I am deluded enough to believe I am the US. All bets are off, it seems, in circumstances in which I suffer from such a bizarre delusion. Plausibly the truth of such biconditionals is far too haphazard to make good sense of my here and now having the relevant weird contingency plan.

One might worry that this argument proves too much. I want to allow that I can, for example, plan what to do if I am in circumstances just like those facing Superman or an android or an intelligent dolphin. The problem is that in the nearest possible worlds in which I believe I am in those circumstances I will also be deluded – in which case the preceding argument has proven too much. At least, it has proven too much for me to have shown that collectives pose a special problem for Gibbard. Moreover, in my view, it would have proven too much in any case since I think it is plausible that we *can* form contingency plans for such weird scenarios. It is, though, worth noting that even if this objection cannot be met that the argument on offer here would *still* provide a challenge to Gibbard's view. The problem would simply not be one unique to judgments about collectives, but would also arise in cases involving radically different agents – perhaps any sort of agent which is such that it would be

metaphysically impossible for the person making the normative judgment to be identical to an agent of that kind. However, in my view the problem *is* unique to collectives.

How, then, do these cases differ from those involving collectives like a nation-state? Here is one relevant difference. There are at least *some epistemically* possible worlds in which I could, without delusion, judge that I am circumstances just like those facing Superman (etc.). To see this, all one has to do is think about radical sceptical scenarios. Drafting this paper right now might be a dream I am having, and I might awaken to find myself in Superman's circumstances, for example, or those of an android or an intelligent dolphin. Granted that, given the facts as we take them to be, it would be metaphysically impossible for me to have been a dolphin, the point is that this is not the relevant form of possibility. Instead we should focus on what is *epistemically possible.* Gibbard is admirably clear about this:

> We might worry that it is metaphysically impossible for me to be Caesar. Since I am not the same person as Caesar, there is no way the world might have been that would have constituted my being Caesar...I answer that…the possibilities we are speaking of are epistemic. I might for all I know for utterly certain, be the genetic son of someone other than my seeming father Harold Gibbard. This, I am fully convinced, is a metaphysical impossibility, since I am his genetic son and the genetic son of anyone else wouldn't be me – but still, it may just conceivably be the case. It is an epistemic possibility, although utterly remote. I can form a contingency plan for this epistemic possibility, deciding whether, in that case, to honour my purported father who raised me or the man who sired me and left. In a like vein, I'm virtually as certain as certain can be that I'm not Julius Caesar, but still, being he is a remote epistemic possibility for which I can plan. (Gibbard 2014: 177)

By contrast, there is no epistemically possible scenario in which I *find myself* in circumstances just like those facing the US.[10] Why? Because there is no "what it is like" for a nation-state like the US. All of the other scenarios I just floated involved rational agents who presumably have some sort of phenomenology. This assumption is essential to the intelligibility of the sceptical scenarios which make it intelligible to suppose that there are epistemologically possible worlds in which I believe *without being delusional* that I am in the relevant circumstances. While I have some idea of what it might be like to "wake from a dream and find myself in the circumstances of Superman facing down Lex Luthor," I have no idea what it could even mean for me to "wake from a dream and find myself in the circumstances of the US on the eve of the Gulf War."

How might this difference help sharpen the challenge for Gibbard? There are two options. We could stick with the biconditional given above, but hold that what counts as "close" when it comes to evaluating possible worlds is context-sensitive, and in particular it is relative to our theoretical interests. In this context, we might argue that one world is closer than another insofar as I am not deluded in that world but I am in the other world. We could instead alter the content of the biconditional so that the relevant possible world is *defined* as one in which the agent *reasonably* believes he or she is in the relevant circumstances. This would also sort the cases appropriately.

Here is a straightforward way of developing this into an argument against Gibbard's view. First, as common sense assumes, it is possible for individuals to make judgments about what a collective agent ought to do in a given set of circumstances. Second, if Gibbard's

[10] This suggests another objection, namely that intention presupposes not being certain you cannot fulfil the intention. If someone is absolutely certain they will never face the relevant contingency (as in the case of an individual planning for being a collective agent), then there is a sense in which they should be certain they will never fulfil the intention, save perhaps in a vacuous sense. In the relevant sense, I do not fulfil my intention to try curry if in London by blowing up London, though that does ensure that I am never both in London and not trying curry. Developing this objection would, however, require more controversial commitments than the arguments in the text.

theory were correct then, necessarily, an individual could make a judgment about what a collective agent in a given set of circumstances ought to do only if the agent could form a contingency plan for what to do if in the circumstances facing the collective. This is an obvious corollary of Gibbard's theory. Third, it is not possible for an individual to form a contingency plan for what to do if in the circumstances facing a collective agent. This is an upshot of the discussion of functionalism above. Therefore, Gibbard's theory is incorrect. The key premise of this argument is the last one (that individuals cannot form contingency plans for being a collective), and I do not pretend to have given a knock-down argument for it. Indeed, I have not even given a knock-down argument that it follows from the most plausible form of functionalism. However, I think I have done enough to make this premise plausible insofar as one is a functionalist, thus putting pressure on Gibbard's view.

Suppose this argument's central premise is false – suppose we can make sense of the very idea of contingency plans by individuals for what to do if they were in the circumstances facing some collective. The idea would be that we just have to accept that such plans are weird, and that having such intentions depends on what one would try to do if one were deluded in various ways. Even granting this much, there would *still* be a powerful challenge for Gibbard's theory in this vicinity. Insofar as we "bite the bullet" and allow for such bizarre intentions, it becomes implausible that there will be a necessary co-extension between people who make judgments about what collectives ought to do and people who have the corresponding contingency plans. I think the US ought to have gone to war with Germany in 1940, but am I such that if I were mad enough to believe I was the US in 1940 I would try to go to war with Germany? Perhaps, but if so that would presumably just be a lucky accident since I would be quite different in many ways if I were mad enough to harbour such a belief, after all. If Gibbard's theory were correct, though, then that theory plus the functionalist account of intentions we are now considering would entail that my normative judgment

entails that I would try to go to war in such scenarios (and that here and now I would believe that I so intend – our first argument above), etc. It certainly at least seems *possible* for me to make a normative judgment about what the US ought to do without such counterfactuals being true of me – and hence without having the needed intention, on the account of such intentions under consideration. The problem is getting two things to *line up* in the right way – (a) when people intuitively make the relevant normative judgments, and (b) when the relevant weird counterfactuals are true of them. It seems unlikely that these two will align, much less align in all possible cases.[11]

We therefore now have the following (third) argument against Gibbard's theory. First, if Gibbard's theory were correct then, necessarily, an individual makes a judgment at time t about what a collective agent in a given set of circumstances ought to do only if the agent at time t has a contingency plan for what to do if in the circumstances facing the collective. Second, necessarily, someone counts at time t as having a contingency plan for what to do if in the circumstances facing a given collective only if he is such that he would try to act as he judges the collective ought to act in C if he believed himself to be in C himself (where C includes all of the features of the collective's situation, including its being the relevant collective). This is just an implication of the theory of such intentions discussed above. Third, it is possible for someone to make a judgment at time t about what a collective agent in a given set of circumstances ought to do without at the same time being such that he would try to act as he judges the collective ought to act in C if he believed himself to be in C

[11] This problem is related in interesting ways to Gibbard's discussion of planning for circumstances in which you know your plan will almost certainly not efficacious, as with the alcoholic who plans not to drink on Saturday when he craves a drink. Here too, the relevant counterfactuals will not be satisfied. See Gibbard, 2003: 51-52. In these cases, though, at least the agent will characteristically satisfy some of the other biconditionals associated with intending – e.g. he will believe that he intends not to drink on Saturday. Also, we would expect someone with such a plan to do things in advance of being tempted to make it less likely that he will succumb to temptation or to avoid temptation altogether. In the case of planning for being a collective, none of these counterfactuals make sense either, which makes the idea that we can plan for such circumstances more obscure. Thanks to an anonymous referee for prompting me to discuss the analogy between the case discussed in the text and these sorts of cases.

himself, where C includes all of the features of the collective's situation, including its being the relevant collective. This is plausible in light of how haphazard the truth of such counterfactuals would be, depending on random facts about what the nearest worlds in which the person is mad in the right way would be like. Therefore, Gibbard's theory is incorrect.

So far, my arguments have all been entirely directed at Gibbard's theory in *Thinking How To Live*. So much the worse for Gibbard's theory, one might respond, but why think this indicates anything about expressivism more generally? I concede that judgments about collectives are problematic only for a certain class of expressivist theories. The problem as I have so far formulated it relies heavily on the functional role of planning/intending. The problem therefore seems not to generalize to paradigmatic expressivists who understand normative judgments in terms of some other desire-like state. For example, A.J. Ayer famously explained moral judgment in particular in terms of feelings or emotions (Ayer 1952). Charles Stevenson's theory was couched in terms of "approval," where this was also understood in terms of emotions (Stevenson 1944). Simon Blackburn emphasizes the idea of a "staircase of emotional ascent," also suggesting an emphasis on emotion rather than intention (Blackburn 1998).

Indeed, Gibbard's own account of specifically *moral* judgment, as developed in *Wise Choices, Apt Feelings*, is immune to the problem, at least with a minor modification. To a first approximation, Gibbard there understood the judgment that someone's action is morally wrong as a judgment that it would be rational for him to feel guilty for doing it and for others to resent him for doing it.[12] He then argued that to make a judgment about when guilt or resentment are rational is to accept a norm which prescribes that others feel resentment towards the agent for performing it and that the culprit feel guilty for having performed it. It is not entirely clear that it makes sense to think of collective agents like the US as "feeling

[12] Allan Gibbard, *Wise Choices, Apt Feelings* (New York, Oxford University Press, 1990, p. 42)

guilty," but we can charitably modify Gibbard's account to take this into account. Instead of insisting that to judge an action is wrong is to deem *both* guilt on the part of the culprit *and* resentment on the part of others rational, we can instead hold that one counts as judging that an action is wrong if one *either* deems guilt on the part of the culprit rational or deems resentment on the part of others rational. Since there is nothing obviously impossible or irrational about resenting the actions of collective agents, this lightly modified version of Gibbard's theory seems to handle judgments about collectives easily enough.[13]

It might now seem like the problem arising from judgments about collective agents is a problem only for a narrow range of expressivist theories. However, the features of Gibbard's theory which make it vulnerable to this problem are not so easily abandoned. In particular, the features of Gibbard's theory which lead to this problem are arguably features which are essential to any adequate account of judgments about *agent-relative* norms. The easiest way to see the connection to agent-relativity is to consider an initially tempting strategy for accommodating judgments about collective agents within a framework that is very similar to Gibbard's in *Thinking How to Live*. Instead of understanding judgments about others as contingency plans for what to do if in their circumstances, we could instead understand such judgments as intentions to do whatever one can to bring it about that anyone in the circumstances facing the person being judged performs the relevant action. This approach avoids any reliance on contingency plans for what to do if one finds oneself in the circumstances of a collective agent of the relevant kind. It thereby neatly avoids the problems arising from the very idea of such plans.

However, any theory with this structure faces another problem. In particular, the proposed approach only works for *agent-neutral* normative judgments, on which a reason for anyone to do something is a reason for anyone else to promote that person's doing it. Given

[13] Thanks to an anonymous referee for pressing me to discuss this point.

agent-neutrality, whenever I judge that someone else ought to do something I will be rationally motivated to promote that person's doing it if I can. Crucially, making an agent-relative normative judgment that someone ought to do something does *not* commit one to thinking that anyone who can promote the person performing the action has reason to do so – even a tiny reason. Agent-relative norms index reasons to individuals. Egoism is a simple and familiar agent-relative norm, but the point applies to any agent-relative judgment about the thing to do. Deontological norms, for example, are very commonly understood as being agent-relative, as the norms corresponding to our "special obligations" to our nearest and dearest, and what Nagel famously called "reasons of autonomy." [14]

Suppose I am playing chess and I judge that my opponent, Jones, ought to move his knight. Plausibly my normative perspective on the reasons germane to our game will be agent-relative – I have reason to do what would promote my winning the game and Jones has reason to do what would promote Jones's winning the game. Does this mean I *thereby* take myself to have reason to promote Jones's moving the knight? Obviously not! It is not as if I would be rationally motivated by my normative judgment to offer Jones some friendly advice. To be sure, I might do that (e.g. to make the game more interesting), but the agent-relative judgment that he ought to move the knight should not as such rationally motivate me to do this – not even a teeny bit. The problem is that the proposed expressivist theory *entails* that I *would* be rationally motivated to tell Jones to move his knight. On that theory, my judgment in this case would be identical to my intending to do whatever I can to promote anyone in Jones's circumstances moving their knight. Since I easily can promote Jones's doing so by pointing out that he needs to move his knight to avoid checkmate, this intention would rationally motivate me to do so. This, though, is perverse.

[14] For a classic discussion of these species of agent-relative norms, see Nagel 1986, chapter 9.

Moreover, this is *precisely* why Gibbard's theory is formulated as it is. Normative judgments need to be understood as contingency plans for what to *do* if I were in the person's circumstances, and *not* as plans to promote anyone's performing the action in the relevant circumstances insofar as I can. Here is Gibbard:

> Suppose I disagree with Caesar on whether, if in his shoes, to go to the Senate on the Ides of March. What sort of disagreement is this? To disagree with his decision I don't need to be against it. Brutus too, after all, may disagree with Caesar on this. For though Brutus plots desperately to ensure that Caesar come to the Senate that morning, he may nonetheless say to himself, 'If I am Caesar in Caesar's shoes this morning, let me stay home'. (Brutus, in technical jargon, thinks Caesar to have agent-centered reason, all told to stay home. At the same time, he thinks himself to have agent-centered reason, all told, to ensure that Caesar leave home and come to the Senate). (Gibbard 2003: 68)

Interestingly, the same objection can be lodged against versions of expressivism which eschew intentions and analyse normative judgment in terms of some other pro-attitude, like desire or preference. Insofar as these pro-attitudes are irreducibly *de se*, they will likely face a version of the problem posed for Gibbard. In that case, the relevant pro-attitudes are still conditional ones concerning what the agent would want if they found themselves in the circumstances of the relevant collective. If, on the other hand, they are not irreducibly *de se*, but are instead impersonal desires to promote anyone's acting in accordance with the relevant norm, they will face the objection that they do not adequately accommodate agent-relativity.

The problem I have laid out here is at least this general, then: Any expressivist theory which aspires to give an account of the nature of judgments about what one ought to do, all things considered, will *either* be unable to account for normative judgments about collectives *or* be unable to provide a plausible account of agent-relative normative judgments. It is

worth noting that it is not very plausible to construe judgments about what one ought to do, all things considered, as judgments about the rationality of some reactive attitude like guilt or resentment. After all, the norms a given agent takes to settle what one ought to do all things considered need not be moral. Since many of the arguments for expressivism in the moral case are comparably good arguments for expressivism about the all things considered 'ought', the dilemma is a pressing one for a wide range of expressivist theories.

II. Objections and Replies

The arguments of section I rely on functionalism about intentions – that is, a functionalist account of what it is to have an intention with a given content. An initially tempting reply is therefore simply to reject functionalism. In fact, it is not entirely clear that abandoning functionalism will help, though. Although I motivated the biconditionals about intention in section I via a functionalist framework, those biconditionals are independently plausible. One might, therefore reasonably still want to endorse those biconditionals even if one rejects functionalism about intentions. In that case, the argument of section I can survive the rejection of functionalism about intentions.

Second, abandoning functionalism about intentions is dialectically problematic for expressivists. A core motivation for expressivism is to avoid the problems facing realists by starting *not* with a metaphysical account of normative facts or states of affairs, but instead with our practice of making normative judgments. In particular, expressivists urge us to notice the distinctively practical *functional role* that these judgments play. On the basis of this functionalist account of normative judgments, the expressivist argues that we should understand them as broadly desire-like states (or, anyway, partly constituted by such states; see section III). Insofar as the expressivist understands normative judgments as reducible to intentions (or "planning states" in Gibbard's sense), they are thereby committed to giving a functionalist account of intentions too. Moreover, the characterization of normative

judgments in functional terms fits best with a broader functionalist approach to the nature of mental states. Expressivists could in principle abandon functionalism. In that case, though it would be unclear how their account of the functional role of normative judgment tells us anything about the nature of normative judgments.

To get an idea of just how central functionalist ideas are to paradigm expressivists, consider Simon Blackburn's *Ruling Passions*, a landmark book in expressivist theorizing. The first chapter is all about the functional role of ethical judgment. Section 2 of that chapter is entitled "Inputs and Outputs" and contrasts the practical functional role of ethical judgments with ordinary factual judgments (Blackburn 1998: chapter one). Here is a representative passage:

> We can usefully compare the ethical agent to a device whose function is to take certain inputs and deliver certain outputs. The *input* to the system is a representation…The *output*, we are saying, is a certain attitude, or pressure on attitudes, or favouring of policies, choices and actions. Such a device is a function from input to output: an ethical sensibility. (Blackburn 1998: 5)

Blackburn goes so far as to claim that quasi-realism just *is* a form of functionalism: "A full-dress title [for my view] might be 'non-descriptive functionalism' or 'practical functionalism'".(Blackburn 1998: 77) Gibbard is equally explicit about this. In his more recent book he *identifies* expressivism as a local form of functionalism:

> We can say that such a strategy [expressivism] amounts to 'functionalism' as the term is used in the philosophy of mind for a theory of mental content.[15]

He adds in the footnote to this passage that, "Jason Stanley pointed out to me that what I had been calling 'expressivism' amounts to functionalism as the term is understood in the philosophy of mind." At least on the face of it, rejecting functionalism is not a very

[15] Allan Gibbard, *Meaning and Normativity* (Oxford, Oxford University Press, 2012), p. 111.

promising strategy for expressivists here. The two most influential expressivists of the latter part of the second half of the twentieth and first part of the twenty-first century consider it a core part of their view.

Another reply is simply to deny that we *really* make judgments about what collectives ought to do. Admittedly we often speak that way, but on this account such claims are always elliptical for claims about what some privileged agent or agents who partly constitute the collective ought to do. For example, perhaps in my toy example of a judgment about whether the US ought to go to war with Iraq, one could maintain that such a judgment is really about what the President or members of Congress ought to do – namely not sign a declaration of war, etc. On the face of it this is rather ad hoc. Why, apart from the pressure it puts on the expressivist theory, should we not take the claims of ordinary people at face value?[16] Granted, we may have an interesting philosophical conversation about whether any 'ought' claims are *true* of collective agents, but the arguments of section I require only the modest claim that people *make* such judgments.

Moreover, in many contexts of utterance there may be no salient individual in a privileged position.[17] Suppose I judge that the US ought to pass a gas tax. *Which* individuals do I think ought to vote in favour of them? All of them? Perhaps, but this does not seem to be part of the semantic content of my judgment. The following conversation is intelligible:

[16] It might be objected that expressivism already essentially involves not taking the claims of ordinary people at face value, in which case this is an odd or even hypocritical complaint from the expressivist. This objection would be forceful against "old school" expressivists like A.J. Ayer who explicitly reject the self-understanding of ordinary speakers. It is, however, much less obviously forceful against quasi-realist expressivists who try to accommodate all of the realist-sounding things ordinary people say, including their claims of mind-independence, normative truth and normative knowledge, in a broadly expressivist framework. Rather than denying that there are normative beliefs, or that people make normative assertions, e.g., the quasi-realist insists that these theses are compatible with their theory when properly understood. Whether quasi-realism, if successful, ultimately implies that the thoughts of ordinary folks (as opposed to moral realist philosophers) are misguided is a difficult question which probably depends on the extent to which ordinary folks are committed to various arguably theory-laden assumptions (e.g. that normative belief has the same "direction of fit" as descriptive belief). I will not try to resolve this question decisively here, but my own view is that insofar as quasi-realism succeeds it will allow us to take the claims of ordinary folks at face value. Thanks to one of the editors of *Mind* for raising this natural concern.

[17] Here I in effect elaborate on the discussion of egoism and the Iraq example in section I above.

Jones: "The US should pass a gas tax."

Smith: "Oh, so you think everybody in Congress should vote in favour of such a tax?"

Jones: "No, actually. My own Senator repeatedly promised not to vote for any new taxes when he was campaigning, and I am a Kantian about promise-keeping. Anyway, all I said was that the US should pass a gas tax – I didn't say that the US should pass such a tax unanimously."

Perhaps the advocate of this line of argument could posit some sort of massive semantic indeterminacy about which agent(s) are being targeted, but again this seems ad hoc.

In fact, agent-relative normative perspectives provide an even deeper objection to this proposal. Suppose I am a thoroughgoing egoist, and I also believe that the interests which agents ought to promote are fixed by their ends. I also take nation-states to be agents with their own ends, so I consistently think that they ought to do whatever promotes those ends. This view certainly *seems* coherent, even if it is not all that plausible. Insofar as this is a possible view, though, the proposal under consideration is doomed. A view like this will make it possible for there to be cases in which an agent can coherently judge that a collective agent ought to perform a given action at time t while at the same time thinking that *none* of the agents who stand in the right sort of privileged position ought to do their part to make it the case that the collective performs the action. The egoist might think that the US ought to pass a gas tax because it is in the national interest – however we fix the goals of the US, these include avoiding catastrophic global warming, let us suppose. At the same time, this egoist denies that the President should vote for a gas tax and also denies that any individual member of Congress should vote for such a tax. Why? Because he believes (perhaps truly) that the overriding end for each of these politicians is to remain popular and electable. Since a gas tax would be extremely unpopular, none of them ought to vote for such a tax by the egoist's lights. This does bring out a weird implication of egoism when applied in such antagonistic

cases, and it may even provide a kind of argument against egoism and other forms of agent-relativity.[18] Regardless of whether anything like that follows from reflection on such cases, such a view does at least seem *possible to hold*. Insofar as it is a possible view, though, the suggestion that seeming judgments about what collectives ought to do are always really judgments about what some privileged individual members of the collective ought to do is hopeless. For here we have an egoist sincerely judging that a collective ought to do something while at the same time judging that none of the relevant individuals ought to do their part to bring it about that the agent performs the relevant action.

Yet another reply would be to understand normative judgments about collectives as plans one adopts *within a fiction*. The idea would be that we somehow imaginatively take up the point of view of the relevant collective agent and within that pretence form a plan about what to do. Note that the view is *not* that I here and now form a *real* plan for what to do when engaging in such a fiction. That plan would, on Gibbard's view, constitute my judgment about what I ought to do if I pretend to be (e.g.) the US in such and such circumstances. Instead, the view trades on the idea that when one engages in some sort of pretence, one in some sense adopts plans within that pretence. On at least some views, though, plans within a pretence are not *really* plans in the same sense that ordinary plans are. If I am playing "Star Wars" with my son then my plan to blow up the Death Star is in one obvious sense not a plan – at least not if we take the fictional plan's content as face value – that is, as being a plan to blow up the Death Star. One can, of course, try to reduce plans within a fiction to plans for what to do when engaging in such a fiction but such views do not seem entirely true to the phenomenology and they also seem overly intellectual in the case of the role playing activities of young children.[19] In any event, this reply will work only if we

[18] I discuss this at length in Ridge 2005. See also Copp 2007.
[19] For an attempt to defend a standard Humean belief-desire account of pretence, see Funkhauser and Spaulding, 2009. For an opposing view which insists on the irreducibility of belief-within-a-fiction and desire-within-a-fiction, see e.g. Velleman 2000.

understand such "plans" as not reducible to ordinary plans for what to do when engaging in a fiction, since in that case they will constitute judgments about what an individual should do when pretending to be a collective, rather than a judgment about what a collective ought to do. The reply should then appeal to "i-plans" - the imaginative counterparts of plans.[20]

This reply has some plausibility. When thinking about what a collective ought to do, it is tempting to anthropomorphize the collective and try to take up its "point of view" even though one at some level realizes it does not really have a point of view in any phenomenological sense. However, understanding normative judgments about collectives in these terms as a way of salvaging an expressivist theory otherwise much like Gibbard's runs into other serious problems. First, the theory makes false predictions in simple cases of pretence. Returning to my previous example, when playing "Star Wars" with my son and pretending to be Darth Vader, I might well i-plan to destroy a planet inhabited by billions of people. Do I thereby count as judging that someone in Vader's circumstances should commit mass murder? Obviously not! Similarly, when playing a war game, I may take up the point of view of the US and i-plan to drop nuclear weapons on Berlin, say. This is consistent with my judging it would be morally wrong ever to use such weapons on civilian targets.[21]

---

[20] Thanks to an anonymous referee for encouraging me to discuss this reply.

[21] In the text I focus on an appeal to imagined planning which has the agent making the judgment imagining him or herself as a collective, but at the same time pretending such collectives have a phenomenological point of view. An alternative version of the appeal to imagined planning would be to imagine oneself *as* an individual but an individual in precisely the circumstances facing the US (e.g.) *apart from* being a collective – in effect, pretending that the US is an individual rather than pretending that it is a collective with a phenomenology. One might then plan for such circumstances in an ordinary sense of 'planning' but against the backdrop of a merely imagined possibility, which would be different from the sort of i-planning discussed in the text.

This is an interesting alternative, but it is not clear to me that it is possible to imagine an individual which is literally *just like* the US in all respects apart from being an individual. An individual, e.g., cannot literally "go to war" – though they might bring it about that the US goes to war (Cf. my discussion of the attempt to understand judgments about collectives as disguised judgments about privileged individuals in the text). One could instead imagine an individual which is "as similar as possible" to the US while still being an individual, but it is far from obvious that such similarities will be sufficient to ensure that what one judges the US ought to do is what such an individual should do. Much more could no doubt be said about all of this, but until a worked out version of the expressivist appeal to planning within a fiction of some kind has been developed it strikes me as premature to go into great detail discussing the potential problems of such a view. Thanks to an anonymous referee and an editor of *Mind* for pressing me on this point.

A second problem is that there will likely be new forms of the "Frege-Geach" problem arising from this approach. After all, on the proposed view normative judgments about ordinary individuals express plans, whereas normative judgments about collectives express mere i-plans. This looks like a problematic shift in meaning between contexts in which individuals are being evaluated and contexts in which collectives are being evaluated.

Third, the expressivist will now need a plausible and independently motivated theory of disagreement which can explain why people who engage in different pretences when engaging in a fictional perspective thereby disagree. It is one thing to agree with Gibbard that when two people plan to act in different and incompatible ways they thereby disagree on the thing to do; that has some independent plausibility. By contrast, it does not seem so plausible to suppose that two people who i-plan to act differently within the same fictional framework disagree. If I i-plan to act one way when pretending I am Darth Vader and you i-plan to act in another (incompatible) way when you are pretending you are Darth Vader, then we simply prefer different forms of role-playing. That should not be sufficient for us to thereby count as having a disagreement.

Having explained why what I take to be the most tempting replies to the arguments of section I do not work given traditional forms of expressivism, I now offer a solution. I argue that a hybrid or "Ecumenical" form of expressivism can do better because of the indirection it affords.

III. Ecumenical Expressivism and Collectives

What is needed is a form of expressivism which does all of the following:

(i) Analyses normative judgments as being at least partly constituted by intentions (thus explaining how normative judgment settles the thing to do),

(ii) does so in a way that smoothly accommodates agent-relativity,

(iii) and is not vulnerable to the arguments of section I (the arguments from functionalism and the impossibility of contingency plans for what to do if a collective).

I argue that a form of "hybrid" or "Ecumenical" expressivism manages to do all three of these things. In fact, I expect a variety of different hybrid forms of expressivist could manage this so long as they all characterized normative judgments as being partly constituted by intentions. I here focus on the version of Ecumenical Expressivism I have defended at length *Impassioned Belief*, not only because it is the version of the approach I favour but because it has been worked out in more detail than any other form of Ecumenical Expressivism. However, it would be interesting to see whether the same moves could not (as I suspect) be made within the structurally similar frameworks of Mark Schroeder's "relational expressivism" and Teemu Toppinen's form of hybrid expressivism.[22]

Very roughly, hybrid theories differ from "pure" forms of expressivism and cognitivism in that their characterization of normative judgment involves both cognitive and conative elements. Not all hybrid theories should be understood as forms of expressivism, though. Some hybrid theories privilege the cognitive element in their theory, providing a meta-normative account of the representational content of normative thoughts. Such theories are still hybrid theories because they incorporate an element of desire into their account. Because this approach in effect begins its account with normative ways the world might be, and explains normative judgments as representing the world as having certain normative features, it cannot avoid many of the standard worries about realism that expressivism was meant to avoid – the open question argument, moral twin earth, arguments from disagreement, as well as various metaphysical and epistemological worries. To be sure, such

[22] See Ridge 2014, Schroeder 2013 and Toppinen, 2013. For an earlier and somewhat different version of Ecumenical Expressivism, see Ridge 2006 and Ridge 2007.

views can do better when it comes to arguments from motivational internalism, but this is only one of the reasons to be an expressivist, and arguably not the most powerful one. Ecumenical Expressivism, by contrast does not appeal to normative contents for representational beliefs at all. Rather, to make a normative judgment an agent must have what I call a "normative perspective" and a representational belief with an ordinary descriptive content, where the contents of these are linked in the right way. Normative perspectives are diachronically stable high level intentions to reject standards of practical reason with certain features and to deliberate and act only in ways that would be allowed by those standards not yet ruled out. The belief component of a normative judgment concerns what any standard of practical reason which is not ruled out (where those not ruled out are "acceptable") would be like in various ways.

Ecumenical Expressivism pretty clearly achieves (i) and (ii) of the three criteria. Normative perspectives just *are* high level intentions, so (i) is trivial. The account has an irreducibly *de se* structure, too, in that normative perspectives always concern how the agent making the judgment will act in various circumstances. Because nothing in the account entails that these standards cannot have an agent-relative form, (ii) is also accommodated.

The crucial question is how Ecumenical Expressivism achieves (iii). It is perhaps simplest to see how it does so by working through an example. Suppose I judge that the US ought to adopt a gas tax. This judgment will be constituted by a normative perspective/belief pair. The normative perspective will rule out a range of standards of practical reason, while the belief will concern what all those standards not ruled out have in common. In particular, the content of the belief will be that all such standards would require the US to adopt a gas tax. Nothing in this theory involves an agent making a judgment planning what to do if he found himself in precisely the circumstances facing the US. His normative perspective rules out a range of standards of practical reason, so he intends not to adopt any of those, and it

also commits him more positively only to act in ways that would be allowed by all of the remaining standards. Once we have this notion of a range of acceptable standards in place, the agent can form beliefs about what those standards would require of other agents – including agents radically different from himself like collectives.

Returning to my example, my normative perspective might rule out all standards of practical reason other than egoism. Crucially the logical form of the egoist standard I deem acceptable might be a universally quantified standard which ranges over *all possible rational agents*. As such, the standard would apply equally to collective rational agents if any there are as it does to individual rational agents. Insofar as the standards I deem acceptable do have this logical form, they entail prescriptions to collective agents like the US. The agent making the judgment can therefore work out what prescriptions they entail for such collective agents without at any stage himself *planning what to do if he were a collective*.

So condition (iii) is also satisfied. The reason a hybrid theory like Ecumenical Expressivism is especially well suited to do this is the *indirection* provided by incorporating both a desire-like element and a belief in the ground-level account of normative judgment. So long as the standards fixed by the agent's normative perspective have universal scope with regards to all possible rational agents, they will trivially include collective rational agents in their scope. The agent can then rationally form beliefs about what these norms entail for various collectives. Of course, *beliefs* about collective rational agents do not raise the worries discussed in section I. There is no argument from functionalism that there could be no such thing as the belief about what e.g. the principle of utility would demand of the US. Because the desire-like element (the normative perspective) does not involve planning what to do if a collective agent, there is no way for the arguments from section I to gain purchase against Ecumenical Expressivism. By "offloading" everything specifically about collectives onto the belief component of the agent's normative judgment, the arguments can be neatly evaded.

This, though raises an important question about the theory's resources. The account works smoothly for any judgment where the norms implicated all range over all possible rational agents. What, though, about someone who judges that there are reasons which can apply only to collective rational agents? For example, someone might think that the fact that performing a given action would carry out the "general will" of the individuals who constitute some collective agent A is itself a reason for A to perform that action. Insofar as this sort of reason necessarily applies only to collective rational agents, the account I have just laid out might seem not to deal very easily with cases like this. Insofar as my normative perspective functions to guide my choices, where I am necessarily an individual agent, why would that perspective discriminate between standards which *necessarily* apply only to collective rational agents?[23]

This is an interesting worry I intend to investigate further in future research. Unfortunately, an adequate discussion of the many issues arising here would go beyond the scope of this paper. However, the worry is also not obviously decisive. One plausible response is to deny that it makes sense to think there could be reasons which necessarily apply *only* to collective rational agents. Insofar as a suitable Kantian universalizability principle according to which any maxim of practical reason must be possible for *all* rational agents to accept is sound, this would provide a principled basis for this view. Of course I do not have the space in this paper to explore the plausibility of such Kantian principles, nor the interesting question of how well they cohere with a broadly expressivist framework.[24]

Won't this response sit poorly with the pre-theoretical intuition that normative judgments which seem to rely on principles which can only apply to collectives *are* coherent and intelligible, though? This is not as obvious as it might seem. Take the example of the reason

[23] See Ridge 2005 for related discussion.
[24] R.M. Hare, of course, famously argued that his form of expressivism fit well with a strong form of universalizability, form which he then tried to derive substantial normative conclusions. See Hare 1952.

to implement the "general will." Plausibly this reason should not be understood as fundamental. Rather, it should perhaps be understood as derived from a more basic normative principle of some kind. Perhaps it can be derived from some version of the categorical imperative. Or perhaps it can be derived from a more basic principle which requires agents to act with integrity, since carrying out the general will of the individuals who *constitute* ones as a (collective) agent is intuitively a form of integrity. Indeed, there is a rich and long tradition of drawing on analogies between such reasons of integrity for individuals and for collectives, most famously going back to Plato's account of justice as the tripartite harmony of the soul in *The Republic*. More contemporary Kantians like Christine Korsgaard have done a lot to explain how this "constitutional model" of agency fits very well with a broadly Kantian action theory.[25] If this is right, and this strategy can be generalized, and the problem of norms which as such apply only to collectives is not so forceful. So long as the relevant norms are not fundamental, but can instead be derived from more basic norms which *can* apply directly to individual, the indirect strategy laid out in this section can still succeed.

Even so, one might worry that these kinds of arguments can establish only that the relevant norms are not *plausibly* understood as fundamental. The worry about Ecumenical Expressivism is that it implies that judgments which endorse such norms as fundamental are simply *impossible*, whereas pre-theoretically such thoughts are possible. In fact, I am not sure we have confident and widely shared intuitions about whether such fine-grained thoughts are possible. It is one thing for a theory to entail that normative judgments about collectives are *as such all impossible*, as I have argued Gibbard's theory and others like it do. It is quite another thing for a theory to entail that a certain proper subset of normative judgments one might have thought could be made about collectives turn out to be metaphysically impossible for some subtle reason. Although this may well still be a cost of

[25] See Korsgaard 2009, especially chapter 7.

the view, it is not so obviously a decisive one. Our intuitions on such fine-grained matters may simply not be reliable. Again, though, much more would need to be said to vindicate this theoretical hunch. For present purposes, I shall rest content with having argued that the residual problem facing Ecumenical Expressivism on this front is, at any rate, much less powerful than the one facing other forms of expressivism.

IV. Conclusion

Expressivists have a problem with collectives, but perhaps not an insoluble one. Insofar as Non-Ecumenical expressivism achieves certain important desiderata, it must offer a problematic account of our judgments about what collective agents ought to do. Gibbard's theory puts this worry into very sharp relief, but any Non-Ecumenical Expressivist theory which can make sense of agent-relativity will face a similar problem. Fortunately, Ecumenical Expressivism can do much better in virtue of the *indirection* afforded by its bifurcated structure. Whether it does well enough, all things considered, depends on delicate issues about the extent to which normative judgments which necessarily apply *only* to collective rational agents are intelligible. I have only scratched the surface of this important further issue, which deserves further attention. However, I hope this article has at least demonstrated that collectives provide a new and interesting problem for expressivists, and that Ecumenical Expressivism can resolve this problem considerably better than Non-Ecumenical Expressivism can.

References

Alwood, A. 2015. 'Should Expressivism Be a Theory at the Level of Meta-Semantics,' *Thought*: 13-22.

Ayer, A. 1952. *Language, Truth and Logic* (New York: Dover).

Barker, S. 2000. 'Is Value Content a Component of Conventional Implicature?' *Analysis*: 268-279.

Blackburn, S. 1998. *Ruling Passions* (Oxford, Oxford University Press).

Boyd, R. 1988. "How to Be a Moral Realist," in Geoff Sayre-McCord, ed., *Essays In Moral Realism* (Cornell University Press, Ithaca): 181-228.

Brandom, R. 1994. *Making It Explicit* (Cambridge, MA: Harvard University Press).

Charlow, N. 2013. 'The Problem with the Frege-Geach Problem,' *Philosophical Studies*: 1-31.

_____. 2015. 'Prospects for an Expressivist Theory of Meaning,' *Philosopher's Imprint*: 1-43.

Chrisman, M. 2013. 'Attitudinal Expressivism and Logical Pragmatism in Metaethics,' in G. Hubbs and D. Lind (eds.), *Pragmatism, Law and Language* (London: Routledge): chapter 6.

_____. 2016. *The Meaning of Ought* (Oxford: Oxford University Press).

Copp, D. 2007. 'The Collective Moral Autonomy Thesis,' *Journal of Social Philosophy:* 369-388.

Funkhauser, E. and Spaulding, S. 2009. 'Imagination and Other Scripts,' *Philosophical Studies:* 291-314.

Gibbard, A. 2003. *Thinking How to Live* (Harvard, Cambridge MA).

_____. 2014. *Meaning and Normativity* (Oxford: Oxford University Press).

Hare, R.M. 1952. *The Language of Morals* (New York: Oxford University Press).

Holton, R. 1993. 'Intention Detecting,' *Philosophical Quarterly*: 298-318.

Horgan, T. and Timmons, M. 1992. 'Troubles on Moral Twin Earth: Queerness Revisited.' *Synthese,* 91: 221-260.

_____. 2006. 'Cognitivist Expressivism', in Horgan and Timmons (eds.) *Metaethics After Moore* (Oxford: Oxford University Press)

Joyce, R. 2010. 'Expressivism, Motivational Internalism and Moore,' in Pigden (ed.), *Hume on Motivation and Virtue* (Palgrave).

Korsgaard, C. 2009. *Self-Constitution* (Oxford, Oxford University Press).

Nagel, T. 1986. *The View From Nowhere* (Oxford, Oxford University Press).

Perez Carballo, A. and Santorio, P. 2016. 'Communication for Expressivists,' *Ethics*: 607-635.

Price, H. 2011. 'Expressivism For Two Voices,' in Knowles and Rydenfelt (eds.), *Pragmatism, Science and Naturalism.* (also: http://sydney.edu.au/time/price/preprints/TwoVoicesNew.pdf).

Ridge, M. 2003. 'Non-Cognitivist Pragmatics and Stevenson's 'Do so as Well'' *Canadian Journal of Philosophy*: 563-574.

_____. 2005. 'Univeralizability for Collective Agents,' *Philosophy and Phenomenological Research*: 34-66.

_____. 2006. 'Ecumenical Expressivism: Finessing Frege.' *Ethics*: 302-336.

_____. 2007. 'Ecumenical Expressivism: The Best of Both Worlds?' in Russ Shafer-Lanau (ed.) *Oxford Studies in Metaethics*, v. 2: 51-76.

_____. 2009. 'Moral Assertion for Expressivists,' *Philosophical Issues*: 182-204.

_____. 2014. *Impassioned Belief* (Oxford: Oxford University Press).

Schroeder, M. 2008. *Being For: Evaluating the Semantic Program of Expressivism* (Oxford: Oxford University Press)

_____. 2010. *Non-Cognitivism in Ethics* (London: Routledge).

_____. Mark Schroeder, 2013. 'Tempered Expressivism,' (in *Oxford Studies in Metaethics: vol 8,* Oxford, Oxford University Press, ed. Russ Shafer-Landau: 283-314).

Sepielli, A. 2012. 'Normative Uncertainty for Non-Cognitivism,' *Philosophical Studies*: 191-2017.

Silk, A. 2013. 'Truth-Conditions and the Meanings of Ethical Terms,' in R. Shafer-Landau (ed.), *Oxford Studies in Meta-ethics* (Oxford: Oxford University Press): chapter 8.

Schneider, B. 2010. 'Expressivism Concerning Epistemic Modals,' *Philosophical Quarterly:*601-615.

Sinclair, N. 2016. 'Reasons, Inescapability and Persuasion,' *Philosophical Studies*, 2016: 1-22.

Stevenson, C. 1944. *Ethics and Language* (New Haven: Yale University Press).

Toppinen, T. 2014. 'Expressivism and Moore's Paradox: A Reply to Woods,' *Journal of Ethics and Social Philosophy* (http://www.jesp.org/articles/download/expressivism-and-moores- paradox.pdf).

_____. 2013. ‘Believing in Expressivism,’ (in Shafer-Landau, ed. *Oxford Studies in Metaethics*, vol. 8: 252-282).

Velleman, D. 2000. *The Possibility of Practical Reason* (New York: Oxford University Press).

Woods, J. 2014. ‘Expressivism and Moore’s Paradox,’ *Philosopher’s Imprint*. (http://www.jesp.org/articles/download/expressivism-and-moores-paradox.pdf).

Yalcin, S. 2012. ‘Bayesian Expressivism,’ *Proceedings of the Aristotelian Society*: 123-160.