

Desire, disagreement, and corporate mental states

Olof Leffler

To cite this article: Olof Leffler (2023): Desire, disagreement, and corporate mental states, *Inquiry*, DOI: [10.1080/0020174X.2023.2203368](https://doi.org/10.1080/0020174X.2023.2203368)

To link to this article: <https://doi.org/10.1080/0020174X.2023.2203368>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 23 May 2023.



Submit your article to this journal [↗](#)



Article views: 109



View related articles [↗](#)



View Crossmark data [↗](#)

Desire, disagreement, and corporate mental states

Olof Leffler 

Department of Philosophy, University of Vienna, Vienna, Austria

ABSTRACT

I argue against group agent realism, or the view that groups have irreducible mental states. If group agents have irreducible mental states, as realists assume, then the best group agent realist explanation of corporate agents features only basic mental states with at most one motivational function each. But the best group agent realist explanation of corporate agents does not feature only basic mental states with at most one motivational function each. So corporate agents lack irreducible mental states. How so? I defend the conditional with an argument from disagreement. On cognitivist approaches to desire, desires function to both motivate and represent the world. Yet such desires are subject to a significant amount of disagreement. Reflection on the folk-psychological properties of desire and belief suggest that this disagreement is better explained by a non-cognitivist approach to desire where they do not have both functions. I then defend the claim that realists are committed to at least some cognitivist motivational states. Using the example of fire brigades, I argue that the best realist explanation of group agents involves mental states with both representational and motivational functions. By modus tollens, corporations then lack irreducible mental states, period.

ARTICLE HISTORY Received 6 February 2023; Accepted 27 March 2023

KEYWORDS Group agent realism; corporate agency; desire; besire

I argue against group agent realism. If corporations have irreducible mental states, then the best group agent realist explanation of corporate agents features only basic mental states with at most one motivational function each. But the best group agent realist explanation of corporate agents does not feature only basic mental states with at most one motivational function each. So corporate agents lack irreducible mental states.

CONTACT Olof Leffler  olof.leffler@univie.ac.at  Department of Philosophy, University of Vienna, Universitätsstraße 7, Room C 0210, 1010 Vienna, Austria

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

I defend the conditional with an argument from disagreement. On cognitivist approaches to desire, desires both motivate and represent the world. But cognitivist desires are subject to a significant amount of troublesome disagreement. Reflection on the folk-psychological properties of desire suggests that this disagreement is better explained by a non-cognitivist approach to desire where they only motivate and do not represent.

Group agent realists then face a problem. Using the example of fire brigades, I argue that if we accept irreducible group mental states, the best explanation of many group agents involves mental states that both represent and motivate. By *modus tollens*, corporations then appear to lack mental states, period.

Here is the paper structure. In section, (1) I explicate the relevant rival non-cognitivist and cognitivist accounts of motivational states. In section (2), I introduce my argument against corporate mental states. In sections (3)–(5), I develop the argument from disagreement which supports its first premise. In section (6), I use the example of fire brigades to finish the argument against corporate mental states. Section (7) replies to objections. Section (8) concludes.

1. Motivational states

I shall argue that group agent realists are committed to taking group agents to have desires that do not exist. To make my case, I shall begin by outlining the two relevant competing accounts of desire.

Let us start with a familiar view. Beliefs and desires are often characterised as distinct. Beliefs represent, desires motivate. This distinction is ordinarily spelled out in terms of their different functional profiles: beliefs represent the world, desires motivate agents to change it. Together they produce action.

We may characterise the different motivational functions by saying that beliefs and desires have different *directions of fit* (if we use that term sufficiently ecumenically). Here, beliefs are usually thought to have a mind-to-world direction of fit: they are true or correct if they fit the world. Desires, however, have a world-to-mind direction of fit, so they are satisfied if the world gets shaped to fit them (normally, though not necessarily, by motivating the agent with the desire). These results are their *success conditions*: a belief is successful if it fits the world, a desire is successful if it is satisfied. More could be said about how to cash out directions of fit – for example, in terms of dispositions or

norms on part of the desiring agents. But our primary concern is to say that beliefs and desires have different motivational functions, and for now the brief characterisation here elucidates these.¹

It is often thought that no mental state has more than one of these motivational functions or directions of fit. On a common view, all beliefs represent and have a mind-to-world direction of fit, all desires motivate and have a world-to-mind direction of fit, and no states have more than one. Hence, beliefs and desires are *modally* separable, so it is always possible to have one without the other, and they are also *causally* separable, so either mental state may cause the other (Sinhababu 2009, 2017; Smith 1994). Call this view non-cognitivism about motivational states, and beliefs and desires interpreted this way ‘non-cognitivist-style’.

Non-cognitivism does not rule out that there are other mental states – for example, intentions, emotions, intuitions, or aliefs – but these are then generally taken to reduce beliefs or desires or to lack intrinsic direction of fit. Hence, non-cognitivism is a thesis about an important subset of basic mental states, or states that do not reduce to other states. This paper focuses on these states.

However, many philosophers deny non-cognitivism, even about basic states. They think that at least some mental states *both* represent and motivate, or have both directions of fit (e.g. Campbell 2018; Gregory 2021; McDowell 1978; Railton 2012; Setiya 2007). Call this cognitivism about motivational states. Without necessarily precluding the existence of some non-cognitivist-style beliefs or desires, the idea is that at least some mental states both motivate and represent. They have both directions of fit: they are successful if they do both.

Cognitivism comes in many flavours. First, there is disagreement about what these states aim to represent: most commonly, there are defences of the guise of the *good* or *reasons* (Gregory 2021). Or maybe they just aim at truth. Second, there is no consensus about how to categorise states with two directions of fit. They have been called ‘desires’ (Railton 2012; Gregory 2021), ‘beliefs’ (Campbell 2018; Gregory 2021), or a third kind of state – so-called ‘besires’ (Setiya 2007). Some even think that desires are like perceptual seeming about the normative (Oddie 2005).

We may, however, ignore these distinctions for now. My argument below will be run on an ontological level, and hence apply to all

¹Some philosophers may nevertheless hesitate to use this terminology: for example, List and Pettit (2011) do not. But the point of the terminology is to characterise the differing motivational functions in more depth. My argument below will generalise to alternative characterizations, so the reader should feel free to switch.

cognitivist views. Among the different terminologies, I do however prefer to think of mental states with both directions of fit as a hypothesis about (a kind of) desire. As it is massively controversial whether desires even exist (Laitinen 2014; Smith 1994), that terminology risks being misleading. Similarly, everyone presumably agrees that some non-cognitivist-style beliefs do not motivate, whereas desires paradigmatically do. Hence, cognitivism appears to be a thesis about desire rather than belief. But the reader is free to switch to her preferred terminology here too.

In sum, then, we face two hypotheses about desire: non-cognitivism, according to which desires have only world-to-mind direction of fit, and cognitivism, according to which (at least some) desires have both directions of fit. What does this have to do with corporations?

2. Against corporate mental states

Realism about group agency says that corporations have mental states that do not reduce to those of the individual member agents in the groups. It has started to become the received view about corporate agency (e.g. Björnsson and Hess 2017; List and Pettit 2011; Tollefsen 2015; cf. French 1979). Realism fits non-cognitivism about motivational states well: realists most commonly attribute beliefs and desires to group agents, where beliefs represent, desires motivate, and they together bring about action. Groups combine their states via their internal points of views or decision-making structures, and then enact them.

This picture is open to different ontological interpretations. Some philosophers are functionalists and take beliefs and desires to be individuated by their roles, yet ontologically real independently of our interpretations. There are many possibilities here: we may individuate states by their relations to sensory input (e.g. responses to information), to other mental states (e.g. how they combine), or to action (e.g. what they produce when combined). Others are interpretivists and take mental states to depend on our interpreting agents as having them: usually by taking up the so-called 'intentional stance' to predict group action using mental states. Or we may have some mixed view: even on interpretivism, we may interpret the states as functionally individuated – there is nothing incompatible here.² For now, we may however ignore these subtleties. My argument targets all realist views that characterise mental states by their motivational functions.

²Though see Strohmaier (2020) for an argument against the mixed view.

More specifically, I argue this:

(1) If corporations have irreducible mental states, then the best group agent realist explanation of corporate agents features only basic mental states with at most one motivational function each.

(2) The best group agent realist explanation of corporate agents does not feature only basic mental states with at most one motivational function each.

(C) Corporations do not have irreducible mental states.

By *modus tollens*, the argument is valid. But is it sound? I shall spend the rest of this paper defending it. I shall defend premise (1) with an argument from disagreement against cognitivist desires. Then I shall argue that, on group agent realism, corporate agents seem committed to at least some cognitivist desires. This supports (2). (C) follows.

3. The argument from disagreement

Again, I shall defend premise (1) in the argument against corporate mental states with an argument from disagreement. The argument is metaphysical, not epistemic: it sets out an important type of disagreement as a fact, and then tries to present the best explanation of that fact. This allows us to focus on relevant hypotheses to discuss (cognitivism vs. non-cognitivism), relevant data by which to adjudicate them (based on our experience of desire), and to make progress on the question of which account of desire is best. Spoiler alert: I shall argue that non-cognitivism is better than cognitivism. Yet later I shall argue that realists are committed to cognitivism.

My argument is inspired by an argument from disagreement developed by J.L. Mackie in the context of defending his famous error theory about morality (1977, ch. 1). For Mackie, the best explanation of disagreement in moral views is that they reflect ways of life rather than moral properties, yielding the error-theoretic conclusion that our moral discourse is systematically false.³ But my argument is this:

(1*) Desires are subject to a significant amount of troublesome disagreement on cognitivist accounts.

(2*) If desires are subject to a significant amount of troublesome disagreement on non-cognitivism accounts, desires are better explained as non-cognitivist

³My argument is, however, much less assuming than his. It is not threatened by criticism about things like the existence of moral properties, how we understand moral knowledge, or the details of moral disagreement (cf. e.g. Rowland 2020).

than cognitivist.

(C*) Desires are better explained as non-cognitivist than cognitivist.

Like the argument against corporate mental states above, this argument is valid, albeit via *modus ponens* rather than *tollens*. I also take it to be sound, for I shall defend its premises in the two sections to follow.⁴ After that, I return to the argument against corporate mental states and defend its second premise.

4. Defence of premise (1*): Archer on temptation

The argument from disagreement has two premises that need to be defended, and hence for the first premise in the argument against corporate mental states to be defended. Premise (1*) reads: 'Desires are subject to a significant amount of troublesome disagreement on cognitivist accounts'. I shall argue that this is so because they stand in tension with many of our ordinary, non-motivational, non-cognitivist-style beliefs. This generates troublesome disagreement that we ought to explain.

Now, just desiring different things need not be odd, even on cognitivism. Alex Gregory has, I think convincingly, argued that desiring different things need not amount to a case of disagreement that would make the cognitivist account implausible. Nevertheless, there *is* a deeper kind of underlying disagreement between motivational states with a mind-to-world direction of fit and non-motivational beliefs with the same direction of fit. This is troublesome.

How so? In response to the objection that his cognitivist account of desire (as belief about normative reasons) might involve too many disagreeing beliefs, Gregory writes:

If desiring [that p] is believing [that you have reason to bring about p], your desire [that \neg p] should be understood as the belief [that you have reason to bring about \neg p]. (...) [T]hese beliefs are perfectly consistent: (...) you might well have reasons to bring about p and competing reasons to bring about \neg p. (Gregory 2021, 22.)

This seems right. But it is still possible that non-motivational beliefs may disagree with our desires when both aim to represent the world (Archer 2020). Such conflict cases are not cases of competing reasons to bring

⁴Of course, the unconvinced reader may plug in some other argument against mental states with dual directions of fit here (e.g. Smith 1994; Laitinen 2014).

about p and $\neg p$. They are, rather, disagreements about whether one has a reason to bring about p at all. We may believe that we have no reason to bring about p , but nevertheless desire that p .

Avery Archer's own examples make this admirably clear. He discusses what he calls *temptation desires*, or experienced desires that go against one's own explicit beliefs about what normative reasons one has.⁵ It seems quite possible to (non-motivationally) believe that something does not give you a normative reason to act, but nevertheless desire it. Archer's examples are a desire to smoke when one believes one has no reason to smoke, and a desire to sabotage a co-worker's project out of resentment while believing that one has no reason to do so.

These desires plausibly occur intrapersonally, such as in Archer's cases. But the type of disagreement they generate generalises to the interpersonal case. If desires truly have a mind-to-world direction of fit, there seems to be massive disagreements between them and other, non-cognitivist-style, beliefs we have regarding which propositions are correct. I desire to have another drink. You don't believe I should have one. You desire to spend all your money on long-termist AI research to ward off the singularity. I don't believe you should. The cases are fruitful and multiply – at present, historically, and across cultures.

Hence, there is massive disagreement between desires and beliefs on cognitivist views of desire. This is so because cognitivist desires clash with non-cognitivist-style beliefs, both intrapersonally and interpersonally. There are lots of attitudes directed at the same propositions which stand in tension with each other, even though some of these attitudes are motivating and some are not. But why would we have desires that stand in tension with other attitudes when they both aim to represent the world with the same direction of fit?

However, might we just not think that we sometimes are wrong, so this disagreement is uninteresting? By analogy with a worry in the debate surrounding Mackie's argument, we may wonder whether this disagreement should be explained at all (Rowland 2020, 28–30). But there are two reasons to think this disagreement indeed is troublesome.

First, any social phenomenon is worth explaining just because it is a datum. But this one also stands out more than many others: on cognitivism, agents have different kinds of attitudes with the same direction of fit aimed at propositions, and these seem to stand in tension as they

⁵For Archer, what makes these troubling for Gregory is that they supposedly are rationally permissible. I am less certain about that, but that question is orthogonal to my argument.

aim at representing the same proposition correctly. This tension is interesting for its own sake.

Second, it is worth giving a philosophical account of any phenomenon that might have philosophical upshots. And there is reason to think that disagreement may have that. Mackie's metaphysical argument from disagreement in favour of his error theory about morality is big if true, and because the different accounts of desires I have discussed above suggest the possibility of an analogous argument, the upshots here might be big if true, too. Hence, we should also be instrumentally interested in plunging deeper into what might be a good explanation of disagreement.

5. Defence of premise (2*): the role of desire

The second premise we need to defend the argument from disagreement, and hence premise (1) of the argument against corporate states, is premise (2*): 'If desires are subject to a significant amount of troublesome disagreement on cognitivist accounts, desires are better explained as non-cognitive than cognitive'. How so?

To answer this question, we need to appeal to relevant evidence. I take that to be evidence that directly bears on the question, which in turn is evidence that consists of familiar philosophical 'data points' that we can arrive at by reflection on our everyday, folk-psychological experiences of desire. These are part of the manifest image of the world which we ordinarily experience. Our experiences of the manifest image generate among the most obvious data points we ought to make sense of in any theory about the world.

By contrast, appealing to things like the plausibility of more grandiose theoretical commitments or empirical evidence based on some pre-defined conception of desire is more likely to stack the deck in favour of views we already are predisposed to accept. Grander theories risk conflicting with data about the experiences of some phenomena in virtue of being based on grander commitments: for example, by our conception of normative reasons, as desires might be thought to ground these (Smith 1994) or to represent external reasons (Gregory 2021; Oddie 2005). And empirical studies already require us to conceptualise what we study: for example, by having determined whether we should count some mental states as beliefs, cognitivist desires, or non-cognitivist desires. Yet the present stage of inquiry is prior to that.

Hence, I shall appeal to some familiar points based on our everyday experience of the role of desire. But I shall leverage them in new ways:

not as counterexamples to cognitivism, but as evidence suggesting what the best explanation is of how desire and belief interact in the face of disagreement. While we are comparing different contingent hypotheses, and hence cannot expect to rule one out as impossible, they indicate that the non-cognitivist account of desire and belief is more plausible than the cognitivist one. Hence, these give a more compelling account of troublesome disagreement.

Here they are:

- *Vividity*

Desires can, famously, be vividly experienced (Sinhababu 2009; 2017, ch. 2). They often come in great outbursts, feeling like they pull us towards their satisfaction – and their non-satisfaction ordinarily hurts whereas their satisfaction is pleasant (Railton 2012, 31–32). Just think of thirst on a warm summer day. Similarly, desires help to draw attention to their content, painting it as attractive or unattractive. A bottle of beer looks more attractive on that summer day than during a hangover. Moreover, their phenomenological intensity and the intensity with which they paint their content easily amplify with imagination. The more you think about that beer on that hot summer day, the more irresistible it seems. And the worse it seems when hungover.

Beliefs are not like this (Sinhababu 2017, 42–43). They are representations that p , and much calmer: they are, ordinarily, such that they respond to the evidence. Outbursts, pain, pleasure, attention or amplification rarely enter the picture. Just because I believe that p and $p \rightarrow q$, I do not tend to be drawn to attend to q , and while it is possible to make a belief conscious (in ‘the mind’s eye’), it does not ordinarily get phenomenologically more intensely felt. Beliefs are calmer than desires. Hence, they are experienced very differently.

These differences in phenomenology between paradigmatic beliefs and desires indicates that we should interpret them differently in the light of disagreement. Disagreement raises the question of which account of desires is best, and then the experienced differences are *prima facie* evidence in favour of a view where beliefs and desires are treated as different kinds of states rather than states with the same motivational function.

A famous reply to appeals to phenomenology is, however, that many desires are not vividly experienced at all. This point seems particularly pertinent in the context of potential corporate desires – few think

corporations have qualia (List 2018). But Hume (1739-40) equally famously replied by introducing the concept of 'calm passions', which do not bear the intense phenomenological hallmarks of paradigmatically experienced desires. Rather, they work unseen in the background of a psyche. In my view, the most plausible hypothesis here is that desires can *range* from being calm to violent. They can be *more or less* vividly occurrent or intense. Their abilities to be dispositional rather than occurrent and to amplify with imagination indicate this. Hence, some desires do not feel much, though they can in principle. And if there are corporate desires, they are almost certainly on the calm side. But, like other desires, they differ from beliefs.

- *Epistemic roles*

Beliefs and desires also appear to play different epistemic roles. Yet these collapse on the cognitivist view: if beliefs and desires have the same direction of fit, they should function similarly in the face of disagreement. But they do not, in at least the two following ways.

First, desires are more recalcitrant than beliefs. Desires often linger in spite of evidence that they are hard to satisfy. Sometimes, the heart just wants what it wants. Hence, when desires conflict with the evidence, or we desire what we do not believe, we do not expect to treat them as representing how the world works. I cannot seriously say 'I want *p*, therefore there is value/reason/truth to *p*' either to myself when faced with contradictory evidence or in an argument when I am trying to work out a disagreement with someone else. What I want may not be valuable, reasonable, or true, and while beliefs also may be false, believing that *p* commits you to believing the truth of *p*.

It may be suggested that beliefs are equally recalcitrant (Gregory 2021, 99). But non-motivational mental states which do not respond to the evidence should not, plausibly, count as beliefs. However else they function, beliefs are made correct if they represent the world, which presumably disposes one to respond to the evidence about their veracity. A representational mental state which does not respond to evidence is rather something else, such as an alief (Gendler 2008). Aliefs may represent the world as being a certain way without responding to evidence. *They* may be recalcitrant like desires. But, for the same reason, they are then likely to have to be downplayed in disagreements.

Second, beyond recalcitrance, there is wishful thinking. Wishful thinking suggests that we cannot argue from our own present desires to

thinking that *p* is true, valuable, or a reason. Desires often make us believe things falsely, which indicates that in disagreements between desires and beliefs, desires do not have the same normative standing as beliefs. Presumably, this is because of their differing motivational functions. Someone who drinks wine too often but is worried about alcoholism is likely to go hunting for evidence suggesting that drinking wine is healthy. But satisfying that desire is unlikely to generate a true belief.

A potential response is that beliefs may lead to wishful thinking too (Gregory 2021, 98–99). And perhaps a non-motivational belief could affect other beliefs wishfully: the belief that one's preferred political party will win the election can make one believe that it will implement its policies when in power. This point is however irrelevant. Desire-based wishful thinking suggests that desires which appear to be nothing like beliefs, such as the desire to justify drinking wine, can affect beliefs. Such desires seem nothing like desires that would function to represent the world.

- *Continuity*

The third relevant property here is the continuity, both experiential and motivational, between 'primal' desires and more complex ones. Primal desires are things like hunger, thirst, and sexual attraction, which plausibly may be given a purely evolutionary explanation. More complex desires are those that are unlikely to be given a purely evolutionary explanation, such as a desire for getting into the grad school which is just about right.

These desires are remarkably similar. They share the properties emphasised in the discussion of vividity. They can come in great outbursts, be connected to pleasure or pain, direct attention, and be more or less vivid depending on what one is imagining. And both primal and complex desires often motivate agents to go to great lengths to take means for their satisfaction. Hence, both primal and complex desires appear to have the same relation to means-beliefs: they often set ends that we try to reach using means-beliefs. Therefore, they have different cognitive roles from beliefs. We are likely to be willing to listen to new evidence about how to best reach our ends, but less inclined to change our ends just because the evidence suggests something else (i.e. our desires can be recalcitrant), and too quick to alter some beliefs because we desire something (i.e. engage in wishful thinking).

This continuity is very different from how rationalistically inclined philosophers approach what I have called primal desires. For example, Gregory (2021, 137) distinguishes between the feeling of primal desires and the desires they ordinarily cause, but thinks that only the feelings, not the desires, are the real phenomena of hunger, thirst, or lust proper. The desires are rather beliefs about reasons based on the feelings. Yet this distinction seems *ad hoc*. We need not introduce it.

- *Normative independence*

The final point I want to emphasise is our apparent ability to desire independently of responding to something external to the desire. In particular, we need not respond to something valuable or some reason to desire, since we may desire what is normatively neutral or bad.⁶ At least since Augustine worried about stealing pears, this alleged independence from what should be has been a contentious issue in debates about the nature of desire (Schmid 2021; cf. Stocker 1979; Velleman 1992). However, I want to put it to new use as bearing on our interpretation of desire and belief when they are subject to disagreement: I will just point out that such variation *prima facie seems* possible in very pedestrian cases. This places an explanatory burden on those who argue that desires have a representational function.

Philosophers sometimes present cases of desiring the neutral or the bad as puzzling and bring in outlandish examples to present them, such as Satan saying ‘Evil, be thou my good!’ (Velleman 1992). But this is unnecessarily complex. I want to watch a poor movie. Why? Because I judge it bad. Or I want to pinch my hand when I write. Why? Because it hurts. Not a lot, and not enough to make me suffer intensely, but a bit. Or, instead of something bad, I can desire something normatively neutral. I find myself randomly scrolling through the news even though I already have read them, so I have no reason to do so and it has no value. But I have time available, so there is also no reason against it, and it is not bad to do it. Why do I do it? Just because I want to.

If desires can motivate independently of representing, it seems odd to ascribe them a representational function. Moreover, this point also enhances several of the points mentioned in the discussions of Vividity, Epistemic Roles, and Continuity. Desires that are independent of the

⁶Moreover, desires need not represent something as true, since we may desire things to be other than they are.

normative can be phenomenally strong or weak, and we are talking about all kinds of desires, whether primal or complex. Yet they can all be recalcitrant and involve wishful thinking. This makes them even more pernicious to appeal to in arguments or as counting for or against various beliefs we have ourselves. They are likely to mislead as they can aim anywhere.

Probably the most familiar response to taking cases of desire that are independent of the normative at face value is to claim that there is something *weakly* normative in them, such as a weak reason (Gregory 2021, 102–103). But even that is positing too much. Normatively independent desires seem, *prima facie*, like data points to handle without positing any kind of reasons. Doing the latter seems *ad hoc* – and this response generalises. Any attempt to explain example cases of desire without value, reasons or truth in terms of other factors seems *ad hoc* because the experience of desires which does not aim at the good seems so pedestrian.

I conclude that desires that appear to motivate while being independent of the normative add up with Vividity, Epistemic Roles, and Continuity to support thinking of desires as having a motivational function only, not a representational one, whereas beliefs only have the latter. It follows that the troublesome disagreement we found in premise (1*) is not so troublesome after all: as desires and beliefs have different directions of fit, they do not stand in tension by aiming at the same propositions while having the same direction of fit.

There is, however, an objection to this defence of premise (2*). The data points I have appealed are still *compatible* with the existence of mental states with both representational functions. Maybe my argument at best says something about some paradigmatic beliefs and desires, whereas some unusual desires still may have the relevant cognitive functions.

However, my argument is an inference to the best explanation (between relevant hypotheses, using relevant evidence) about a contingent matter. True, cognitivist desires *could* still exist, but the argument shows that it is *implausible* that they do. In particular, Vividity, Epistemic Roles, and Normative Independence provide *prima facie* evidence against cognitivism. Beliefs and desires look different, they serve different epistemic roles, and they appear able to motivate without representing. And Continuity indicates that desires are a unified kind, so it is implausible to think of some as cognitivist and some as non-cognitivist. (And Normative Independence enhances all the other points.)

In fact, more generally, even if we do not want to run my argument above as an inference to the best explanation but as something else (for example, as a reflection on our phenomenology), it is too demanding to require an argument about contingent matters to show that alternative mental states could not possibly exist. The world could work in many ways. But desires are still more likely to be non-cognitivist than cognitivist. Hence (2*) is defended. As premise (1*) previously has been defended, (C*) follows. Then premise (1) in the argument against corporate mental states is defended, too.

6. Against corporate mental states: premise (2)

I have argued that the best explanation of disagreement between non-cognitivist-style beliefs and cognitivist-style desires is that desires do not have both directions of fit. They are non-cognitivist. It is more plausible to take all beliefs to be representational only and all desires motivational only than to think that some desires represent.

This defence supports premise (1) in the argument against corporate mental states: 'If corporations have irreducible mental states, then the best group agent realist explanation of corporate agents features only basic mental states with at most one motivational function each'. It suggests that there are no basic mental states with more than one motivational function, so if corporations have mental states, these are the states they have.

But what about premise (2)? It reads: 'The best group agent realist explanation of corporate agents does not feature only basic mental states with at most one motivational function each'. How come? I shall argue that group agent realism is committed to the existence of mental states with both directions of fit. *Modus tollens* then suggests that corporations lack irreducible mental states.

For that, let us return to group agent realism. For now, assume that it is true (based on any argument for it that we may prefer). I suggested in section (2) that group agent realists ordinarily characterise group beliefs and desires functionally. Their respective functions are to represent and to motivate, and they link up to produce action. On the Humean theory of motivation, in particular, agents form instrumental desires based on desires for ends and means-beliefs that guide them. The group agent is particularly likely to do so via its decision-making structure or internal point of view. Of course, these can look very different, and there can be many types of corporate agents that differ in virtue of a multitude of

properties: different organisations, capacities, aims, degrees of consensus on its goals among members, and so on. But we may abstract from such possibilities.

Instead, let us consider a real-life case of a group agent: a fire brigade. Fire brigades may no doubt function in many ways, have different capacities, aims, degrees of consensus, and much else besides. Yet any fire brigade has aims ('Quench fires in the area!', 'Rescue cats in trees!') and often receives information that makes it act on its aims ('There is a fire alarm in a local school!', 'A cat in a tree is meowing loudly!'). And it has an internal decision-making structure that outlines how it may go about using information to quench fires: for example, some firefighters will drive firetrucks, others will not. Hence, it is a paradigmatic group agent on the realist view. It is a structured group that will perform intentional actions based on its motivations (e.g. 'Quench fires in the area!') and representations suggesting means to satisfy them (e.g. 'There is a fire alarm in a local school!'). Moreover, on realism, these mental states are irreducible to the states of the member agents.

It is, in fact, *particularly* characteristic of fire brigades to be disposed to quench fires once they receive information about them via their alarms. This is plausibly even part of what differentiates a fire brigade from other group agents: it has these particular aims and responds to information in this particular way. (The clue is in the name.) Hence, it is plausibly a constitutive aim of a fire brigade to eliminate fires when it receives information about them. If the (members of the) fire brigade were to hold a vote and decide to adopt a new aim, so the organisation also were to change its aim – to run for political office rather than quench fires, say – it is plausibly no longer a fire brigade but rather a political party. This makes quenching fires a *central* aim of the fire brigade, where a central aim is the type of aim that consistently motivates it and structures the action projects it engages in. Other aims, such as saving cats, are likely to be less central.⁷

The hypothesis, then, is this: if we are group agent realists who think groups have functionalist-style mental states, then not least because of its centrality, the fire-quenching aim of the fire brigade is best explained as a mental state with both directions of fit. While not all functions need to be mental states – the function of a house may just be to provide shelter – the fire brigade case is based on an organised group, which

⁷It is also possible that the cat-saving aim could be constitutive, central, or both, so it should be analysed like the fire-quenching aim. Alternatively, there could be group agents without constitutive or central aims. But I focus on fire-quenching for simplicity.

realists think has mental states. Moreover, it is set up to respond to information *directly*. Its receiving information also motivates it. This means that we ought to interpret it at least in part as having a desire. But the desire to quench fires seems to have both a mind-to-world and world-to-mind function simultaneously. Hence, its motivational function does not appear in principle separable from its representational function. It looks like, if it is to be explained as having irreducible mental states, it has a mental state which simultaneously responds to evidence and motivates. This is a cognitivist desire.

I will defend this cognitivist account of the desire of the fire brigade with an argument from elimination. The cognitivist explanation appears to be the best among plausible competing explanations of how fire brigades work, as long as we assume group agent realism. This is different from giving the best possible interpretation under the assumption of some particular realist view: on a narrow construal of interpretivism, for example, we may only use non-cognitivist-style beliefs and desires in our explanations. But that begs the question: if we are functionalists of any kind, we should opt for the best functionalist explanation. And then we get committed to a desire with both directions of fit. Why? Let us consider some alternatives.

A first alternative explanation is causal. When the alarm sounds, maybe the fire brigade forms a belief which *causes* a desire, giving it a non-cognitivist belief/desire-pair. And this causal account can be beefed up. Perhaps fire brigades form desires to quench fires that are *reliably* caused by alarms, or are caused under *normal conditions*, or under some other more wide-ranging set of conditions. These alternatives look like more plausible accounts than the first, simple, causal view of the cognitivist desire of the fire brigade, as fire-quenching is central to its motivations.

Yet they all fail to account for the immediacy of the relationship between information retrieval and motivation that fire brigades plausibly have. All causal views allow beliefs and desires to be modally separated: it is possible to have one without the other. But if they are modally separable, as non-cognitivist-style beliefs and desires are, it is possible for the fire brigade to receive the information without forming desires. Yet the connection between information retrieval and motivation is plausibly tighter than that.

The case of malfunctioning brings out this point. On any causal interpretation, one possible malfunction of information retrieval by the fire brigade is that it does not get motivated at all by forming a belief

about the information, for it is possible to believe there is a fire nearby without a desire to quench the fire being caused. And even on beefed-up causal accounts such as the reliable or normal conditions accounts, this could happen over a vast range of possible worlds: worlds that only differ in terms of whether some apples are red or green, say. But if the alarm sounds and an alleged fire brigade does not always immediately come to desire to quench fires over this vast range of possible worlds, it is hard to make sense of the centrality of the motivating desire for the brigade. For a desire that is central to an entity's psychology will motivate them consistently and in a way which structures their action projects: it is not just a set of separate desire tokens that may or – in many worlds – may not be caused by information retrieval. Hence, the causal theories of desire formation allow for *causal gaps* that stand in tension with the centrality of key corporate desires.

On a second alternative explanation, the fire brigade may have a standing long-term desire to quench fires. Then what happens when it receives information from the alarm is that the desire gets an outlet: the brigade acquires a means-belief about how to satisfy its desire, and it can then go satisfy it. A standing desire to quench fires could plausibly be a central or constitutive aim of fire brigades.

Nevertheless, there is something irking about modal separability here too. The standing long-term desire view runs into similar problems as the causal view in the case of malfunctioning. Here, the problem is not that the fire brigade may fail to form desires over a vast range of possible worlds, yielding causal gaps, but the possibility of *motivational gaps*. Over a vast range of possible worlds, it is possible that the fire brigade may fail to have its belief linked up with the desire to quench fires, hence motivating them. Yet for a desire to be central to an entity, like the desire to quench fires is for fire brigades, it should not be possible to have such motivational gaps rather than be consistently motivated. Central desires motivate consistently and guide projects, but they need not always do that if motivation is gappy. Hence, the long-lasting desire view fails to explain the centrality of the fire brigade's desire.⁸ Of course, the motivational force of cognitivist desires may also sometimes be too low to generate action. But cognitivism still allows for some level of motivation to remain even in such cases.

⁸This worry also rules out interpreting it as stemming from other mental states, such as emotions (fear?) or long-lasting intentions (plans?) that might cause or activate desires. All such states are separate from desires if they cause or activate them, which makes modal separability and hence motivational gaps into issues.

A third possibility is to appeal to non-basic, composite, mental states. I have deliberately focused on basic mental states. Yet on a possible non-cognitivist view of intention, intentions consist of belief/desire-pairs. We could interpret them as composite mental states with both directions of fit. Perhaps a fire brigade could have a composite intention based on the information from the alarm plus a desire to quench fires?

This is a plausible interpretation of the fire brigade when it acts, and it does not stand in tension with non-cognitivism. But a fire brigade cannot have a standing composite intention to quench fires when it does not believe there are any fires, viz. has not received any information via the alarm. Then it would only desire to quench fires, but not believe there is a fire to be quenched. It would therefore not have a composite intention prior to receiving information. This takes us back to the motivational gap worry from above. It should not be possible for it to receive information without becoming motivated, but it may do that if its basic mental states are modally separable. For, if so, they exist independently and may fail to link up.

Finally, a fourth possibility is that we still could interpret the fire brigade as having some beliefs and desires, just not a cognitivist desire. It could still be interpreted as having *some* functionalist-style mental states, even though its constitutive aim does not feature or give rise to mental states. Rather, its mental states would consist of other representations and motivations.

However, if what the fire brigade does should be interpreted in terms of mental states leading to actions, its receiving information via the sounding of the alarm and its subsequent motivated actions to respond to it should be interpreted as stemming from one mental state. For this is a paradigmatic case of an action that functionalist-style realists should treat in terms of mental states. There is a problem here because it is a critical test case. And it is a problem not just about fire brigades, but for *any* attributions of functionalist mental states: if they fail with the paradigmatic motivation of the fire brigade, the view does poorly in general.

Having considered these alternative possibilities, the cognitivist explanation of the fire brigade case should start to seem plausible. This gives us premise (2): 'The best group agent realist explanation of corporate agents does not feature only basic mental states with at most one motivational function each'. This gives us (C): 'Corporations do not have irreducible mental states'.

7. Objections to (C) and replies

I have inferred (C) from (1) and (2) by *modus tollens*. But why not run the fire brigade mental state with dual directions of fit as a *reductio* of the argument from disagreement instead? It indicates that there are desires with both directions of fit. Maybe there are such states, after all?

The fire brigade case does however not provide a relevant datum about desires. In the argument from disagreement, I argued for non-cognitivism by reflecting on everyday folk-psychological desires. These are part of the manifest image as we experience the world. The desires attributed to corporations are however based on contentious theoretical assumptions: both functionalism and interpretivism in the philosophy of mind, as well as their extension to corporate agents, have been debated for decades. The cognitivist desires are desires we should attribute to corporations *if* we are functionalist-style realists about corporate agency. But theoretically heavy-duty interpretations do not by themselves count as data to be explained: they may well stack the decks, mislead, or have to be abandoned.

Admittedly, when discussing Vividity, I claimed that corporate desires are likely to be calm. Could it not be the case that corporate agents have some desires, albeit unusual ones? And perhaps they are even more unusual than just not being experienced: perhaps they have dual functions? However, the same point applies here as when I admitted that it is possible for cognitivist desires to exist, but that it is implausible that they do: that is not *the best* explanation of the data. We should prefer an account of corporate agency which fits experience rather than theory in the first place.

In response, group agent realists may claim that corporate desires *do* fit the manifest image. Perhaps this is because we often mention them in ordinary discourse – which we, indeed, do. Yet everyday language use is different from experience. It is plausible that much such talk is metaphorical: atheists can swear saying ‘go to hell’ without seriously believing in hell.

Another possibility is that our interpreting corporations as having desires is enough to make it plausible that they have them. And it is always possible to *interpret* a psychology in belief-desire terms. Could we then not say that corporate mental states are part of the manifest image? In fact, maybe we even should take corporate mental states as we interpret them to involve non-cognitivist style desires and beliefs, given the argument from disagreement?

I agree that the interpretation is possible. But it is unsatisfying for reasons outlined above: the fire brigade case should be explained as it is

a paradigmatic case of action by an organized group, but it does not fit non-cognitivism. This rather suggests that we should do away with corporate mental states in general. If so, we had admittedly better also explain how corporations work without mental states and what goes wrong with arguments in favour of corporate agents. But that is a task for elsewhere.

8. Conclusion

In section (1), I introduced non-cognitivism and cognitivism about motivational states. In section (2), I introduced an argument against irreducible corporate mental states. In the ensuing sections (3)–(5), I then defended its first premise with an argument from disagreement. In sections (6) and (7), I returned to the argument against corporate mental states and defended its second premise and conclusion. As the best realist explanation of corporate mental states commits realists to desires with both directions of fit, we acquire a *modus tollens* argument against group agent realism.

Acknowledgements

I want to thank my colleagues in the ERC project ‘The Normative and Moral Foundations of Group Agency’ as well as an audience at the Mind and Society seminar series, both at Vienna, for discussion of previous versions of this paper. I also want to thank an anonymous reviewer for this journal for extensive feedback. Research for this article has been funded by the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 740922, ERC Advanced Grant ‘The Normative and Moral Foundations of Group Agency’, University of Vienna.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The research for this paper has been conducted in the project ‘The Normative and Moral Foundations of Group Agency’ at the University of Vienna. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme [grant number 740922].

ORCID

Olof Leffler  <http://orcid.org/0000-0001-9073-2552>

References

- Archer, A. 2020. "Are Desires Beliefs About Normative Reasons?" *Analytic Philosophy* 61: 236–251. doi:[10.1111/phib.12169](https://doi.org/10.1111/phib.12169).
- Björnsson, G., and K. Hess. 2017. "Corporate Crocodile Tears? On the Reactive Attitudes of Corporate Agents." *Philosophy and Phenomenological Research* 94: 273–298. doi:[10.1111/phpr.12260](https://doi.org/10.1111/phpr.12260).
- Campbell, D. I. 2018. "Doxastic Desire and Attitudinal Monism." *Synthese* 195: 1139–1161. doi:[10.1007/s11229-016-1255-1](https://doi.org/10.1007/s11229-016-1255-1).
- French, P. A. 1979. "The Corporation as a Moral Person." *American Philosophical Quarterly* 16: 207–215.
- Gendler, T. S. 2008. "Alief and Belief." *Journal of Philosophy* 105: 634–663. doi:[10.5840/jphil20081051025](https://doi.org/10.5840/jphil20081051025).
- Gregory, A. 2021. *Desire as Belief: A Study of Desire, Motivation, and Rationality*. Oxford: Oxford University Press.
- Hume, D. 1739–40. *A Treatise of Human Nature*. Edited by L.A. Selby-Bigge and P.H. Nidditch. Oxford: Oxford University Press.
- Laitinen, A. 2014. "Against Representations with Two Directions of Fit." *Phenomenology and the Cognitive Sciences* 13: 179–199. doi:[10.1007/s11097-013-9328-9](https://doi.org/10.1007/s11097-013-9328-9).
- List, C. 2018. "What is it Like to Be a Group Agent?" *Noûs* 52: 295–319. doi:[10.1111/nous.12162](https://doi.org/10.1111/nous.12162).
- List, C., and P. Pettit. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.
- Mackie, J. L. 1977. *Ethics: Inventing Right and Wrong*. New York: Viking Press.
- McDowell, J. 1978. "Are Moral Requirements Hypothetical Imperatives?" *Aristotelian Society Supplementary Volume* 52: 13–42. doi:[10.1093/aristoteliansupp/52.1.13](https://doi.org/10.1093/aristoteliansupp/52.1.13).
- Oddie, G. 2005. *Value, Reality, and Desire*. Oxford: Oxford University Press.
- Railton, P. 2012. "That Obscure Object, Desire." *Proceedings and Addresses of the American Philosophical Association* 86: 22–46.
- Rowland, R. 2020. *Moral Disagreement*. London: Routledge.
- Schmid, H. B. 2021. *Evil in Joint Action: The Ethics of Hate and the Sociology of Original Sin*. London: Routledge.
- Setiya, K. 2007. *Reasons Without Rationalism*. Princeton: Princeton University Press.
- Sinhababu, N. 2009. "The Humean Theory of Motivation Reformulated and Defended." *The Philosophical Review* 118: 465–500. doi:[10.1215/00318108-2009-015](https://doi.org/10.1215/00318108-2009-015).
- Sinhababu, N. 2017. *Human Nature: How Desire Explains Action, Thought, and Feeling*. Oxford: Oxford University Press.
- Smith, M. 1994. *The Moral Problem*. Oxford: Blackwell Publishing.
- Stocker, M. 1979. "Desiring the Bad: An Essay in Moral Psychology." *The Journal of Philosophy* 76: 738–753. doi:[10.2307/2025856](https://doi.org/10.2307/2025856).
- Strohmaier, D. 2020. "Two Theories of Group Agency." *Philosophical Studies* 177: 1901–1918. doi:[10.1007/s11098-019-01290-4](https://doi.org/10.1007/s11098-019-01290-4).
- Tollefsen, D. P. 2015. *Groups as Agents*. Hoboken: Wiley.
- Velleman, J. D. 1992. "The Guise of the Good." *Noûs* 26: 3–26. doi:[10.2307/2215684](https://doi.org/10.2307/2215684).