

Agent-Switching, Plight Inescapability, and Corporate Agency

ABSTRACT: Realists about corporate agency, according to whom corporate agents may have aims above and beyond those of the individuals who make them up, think that individual agents may *switch* between participating in individual and corporate agency. My aim is, however, to argue that the inescapability of individual agency spells out a difficulty for this kind of switching – and, therefore, for realism about corporate agency. To do so, I develop Korsgaard's notion of plight inescapability, which on my take suggests that individual agents are continuously faced with fully exercising their individual agency (absent external limits at the time of its exercise). But then individual agents may not switch to acting as members of corporate agents if the aims of the corporation differ from their own, in the sense that they have different mental state tokens. But as it is possible to participate fully in the action of a corporate entity, this incompatibility between individual and corporate aims suggests that we ought to think of corporate agents as in some sense reducible to individual aims.

Group agent realism has started to become the received view about corporate agency: or, in other words, the agency of organized entities such as businesses, universities, and states. Realists come in different stripes, but they all think corporate agents have their own psychologies and perform their own actions which are over and above of, but still dependent on, appropriately organized individual member agents' psychologies and actions (Björnsson & Hess, 2017; French, 1979; Hess, 2020; List & Pettit, 2011; Rovane, 1997; Tollefsen, 2015). Group agent realism is descriptively ambitious, and it also has many normative motivations and implications. Typically, defenders take it to lend corporate groups enough agency to be fitting targets of moral responsibility judgements about all kinds of issues, ranging from corruption to climate change – and beyond. But it also provides the ground for a picture of group mental states with implications for epistemic norms (Brown, 2022; cf. Mitova, 2022) and allows us to understand the nature of politically salient actors who we may want to both respond to appropriately and keep in line (Pettit, 2023).

My main aim in this paper is, however, to present a challenge for realism. I do so descriptively: I am interested in the metaphysics of corporate agency. The answer to the challenge may have all kinds of normative upshots, for example about how to treat group responsibility, group epistemology, and group agents politically. But these are downstream from this paper: for now, I focus on metaphysics.

The challenge is this. On group agent realism, the individual member agents who make up corporate agents sometimes act individually and sometimes as part of the larger corporate agents. In the latter case, they enact the aims or desires of the corporation. But to do so, they must be able to *switch* between these types of agency and action at will (List & Pettit, 2011). However, many

constitutivists simultaneously argue that the normative commitments of individual agency are *inescapable* (e.g. Ferrero, 2019; Korsgaard, 2009; Velleman, 2009). These normative commitments are (undefeated) norms and aims that are implicit in someone's agency, either directly or indirectly. In the direct case, they may be partially constitutive of agency itself, such as the categorical and hypothetical imperatives according to Kantian constitutivists.¹ In the indirect case, they are norms and aims that are within the capacity of an agent to enact, such as normative reasons, since the capacity to act on these also is a feature of agency. In either case, the normative commitments are inescapable.

Inspired in particular by Korsgaard, I shall argue more specifically that individual agency and its normative commitments are *plight inescapable*. Explicating this notion, I argue that plight inescapability says that individual agents continuously are faced with exercising their own agency fully absent external limits at the time of its exercise. Hence, they are continuously in a position to enact the normative commitments of their own agency, unless something external to agency at the point in time of acting is in the way. Understood as such, plight inescapability rules out switching into enacting forms of agency – including corporate agency – where corporate agents have irreducible mental state tokens that differ from those of individuals. We see this most clearly when individual and corporate agents have different normative commitments.

This result does not just say something about the psychology of agent-switching: it also generates a challenge for realism. If corporate agents are not treated like separate entities with their own psychologies but instead reductively, and hence without irreducible mental states, individuals need not switch when acting collectively. If so, what they do when acting in corporate contexts may be understood in terms of their own individual commitments instead. As corporate agency with full agential participation is possible, the argument against agent-switching rather supports reductionism.

Here is the paper structure. In section (1), I explicate the nature of agent-switching. In section (2), I introduce the argument against it. In section (3), I defend its first premise, viz. that individual agency is plight inescapable. In section (4), I defend its second premise by developing an account of plight inescapability in terms of agents continuously exercising their own agency fully (absent external limits at the time of its exercise). In section (5), I defend its third premise by outlining the conflict between exercising one's agency fully individually and as part of a corporation. This

¹ Constitutive aims are desires or drives, like Velleman's self-understanding motive (Velleman, 2009). Constitutive norms are, rather, standards like the imperatives (Korsgaard, 2009). See Ferrero (2019) for the distinction.

Note that I also have added the 'undefeated' qualification because I do not want to count some obviously normatively subpar norms and aims as relevant.

generates the conclusion. In section (6), I show how our being incapable of agent-switching is a problem for realism. I conclude in section (7).

(1) Agent-Switching?

Let us start with the question of how agent-switching works. It can occur on any view where corporate agents are such that they have ends over and above those of the individuals who constitute them. For illustrative purpose, I shall focus in particular on List & Pettit's (2011) account of agent-switching to explain what it is.

Again, it has become popular to think that corporate agents have mental states that depend on, but nevertheless do not ontologically reduce to, the mental states, actions, and interactions of individual agents. The view that group agents can be treated as having such mental states, or at least functional analogues thereof, has a reasonably long pedigree (French, 1979). But it has recently become close to a received view among social ontologists (e.g. Björnsson & Hess, 2017; Hess, 2020; List & Pettit, 2011; Rovane, 1997, 2014; Tollefsen, 2015).² *Realism* about group agency is the view that corporate agents have irreducible mental states that they can act on. Defenders typically claim that groups have motivating desires, representational beliefs, and some capacity to bring these together into belief/desire-pairs on which they may act from their own points of view. But these states and actions do not reduce to those of member agents.³

Group agents nevertheless bear an intimate connection to their members. Their mental states and actions plausibly depend on those of their members without reducing to those of their members. Here, I want to particularly emphasize the way in which the members must be able to *switch* between acting in their capacity as members of the corporate agent they make up and as individuals. This is not just the case for CEOs or presidents but true for all individual members of corporate agents: at the very least, they must be able to carry out their specific roles, or else the corporation would not be able to act on those of its desires that are associated with those roles. For example, assume that my university is made up, in part, by me. Then I can sometimes perform the actions of my university, for example by teaching or by attending meetings, which the university desires to do. But sometimes I can perform the action I desire to do, such as when I go to have lunch at a restaurant instead of buying a meal deal from the supermarket. The possibility of switching between enacting corporate desires and my own is a necessary commitment of all forms of realism.

² It is tempting to cite Gilbert (2009) and Tuomela (2013) here too, and with minor variations, their views fit the pattern. I show in section (6) below how my argument applies to their views too.

³ There is some controversy about how to interpret this picture in more detail – for example, as functionalist, interpretationist, or both (Strohmaier, 2020). But my argument applies to all realist views.

However, List & Pettit think that performing the actions of the group does not just involve furthering its aims based on one's own interpretation of them. Rather, doing so requires identifying with a group: to see oneself as a member of it and as such identify with its attitudes rather than those one would have for oneself independently of it.⁴ This leads to an automatization of one's actions on part of it. Thus:

[T]he members of a group may connect themselves directly to the attitudes of the group. They do not treat the group attitudes as mere indicators of what the group is to do, asking themselves explicitly whether they wish to identify with the group, and acting only if they have this wish. Rather, their individual attitudes are under the automatic guidance of the group, so that they can respond (...) spontaneously (...). Or at least they may do this when there are no 'red lights' that suggest they should hesitate and take stock. If we reach an alignment between what the group requires of us and our individual attitudes through adopting the group's viewpoint, and not merely through an incentive structure under which individual and group interests coincide, then we, the members, each have attitudes in whose propositional expression the group figures as we. (List & Pettit, 2011, p. 192)

Hence, List & Pettit think that individual agents need not interact with irreducible group mental states, such as beliefs and desires, only as if they were external to their own mental states. Rather, individuals can switch into having group mental state tokens, at least *ceteris paribus*. Individual agents can hold and enact group propositional attitudes which are not the individual agents' own, but rather those of a group mind. These are the attitudes 'in whose propositional expression the group figures as we.'

But how may we switch? List & Pettit think that individual agents choose whether to take up some particular identity, whether as individuals or group members, and hence its associated mental states. Here, individual agents are special for three reasons (List & Pettit, 2011, pp. 198-199). First, an agent's identity *qua* individual is *comprehensive*, meaning that the individual can have attitudes about everything. Second, it is *inalienable*, meaning that individuals cannot opt to implement group identities and their associated aims only. Third, it is *proprietary*, meaning that it is not necessarily subject to the control of other agents like group identities are. Individuals may then switch into

⁴ List & Pettit add that they think that this distinction has much explanatory power. First, they take identification with certain beliefs and desires of a group to be a pre-requisite for acting on them. Second, they think identification with corporate agents allows them to perform actions as persons, and therefore 'enter a system of obligations, making claims on others, authorizing their claims on me, and doing all this as a matter of common awareness' (List & Pettit, 2011, p. 178). Third, identification allows groups to come together, therefore allowing them to 'corporatize' and act as responsible agents. Fourth, identification with corporate agents may explain how agents treat their interests as important – sometimes too important, comparable to Rousseauvian amour-propre.

identifying as group members while maintaining an underlying individuality which is comprehensive, inalienable, and proprietary.

Nevertheless, individuals performing group actions need not always act in their own interests or for the reason that something supports their own aims. They may be so emotionally entangled with some group that they do not just identify with it, but happily sacrifice their own aims to those of the group. Importantly, an individual who identifies with a group may then act on aims set by group-level aims or desires rather than their own: 'rationality at the group level may require [individuals] to act on the basis of judgments and preferences that they do not hold themselves, and that are not necessarily upheld by a majority – in some cases, not even by a minority – of members' (List & Pettit, 2011, p. 199).⁵ A corporation may desire to raise prices even though an individual employee may not, and an individual identifying with the group may then identify with and enact the group desire rather than their own.

To identify with a group in this way is thus to switch into acting as part of a corporate agent. It is fruitful to elucidate List & Pettit's account of identification using the terminology of another kind of agent-switching recently developed by Thi Nguyen (2020). Nguyen aims to explain how we can engage in striving game-play – a special type of aesthetic agency where we aim to win in the face of obstacles specified by the rules of the games we play, even though these may seem to impede our deeper aims – but his terminology has also been generalized to apply in other domains by Ferrero (2021) and Kukla (2021). As Nguyen observes, it seems like we may switch between our ordinary aims and the aims of games. And as Ferrero and Kukla observe in slightly different ways, and I shall exemplify below, we do seem to switch between various types of agency in life quite often. This will help to elucidate agent-switching.

For Nguyen, agents come in layers. There is an outer layer consisting of our personal long-term aims. But there are also inner layers constituted by sub-aims that we may take up for the sake of reaching the outer aims. These can have all kinds of features, including of being such that they seem to go *against* some of the outer aims. This is nothing strange – we can easily constrain our aims, whether by tying ourselves to masts or by following the rules of a game like monopoly – so Nguyen thinks games designers make use of our ability to take on various kinds of constraints to refine inner layers for aesthetic purposes. The types of agency we take on allow us to engage in immersive activities in the inner layers even when this, seemingly, might seem to stand in tension with our outer-layer aims. Our outer-layer aims may involve making loved ones happy, not to

⁵ Why would this be a feature of *rationality* on the group level? The key point is that because group mental states are not reducible to individual ones, rational group attitudes need not be the same as individually rational ones either (cf. List & Pettit, 2011, ch. 2-3).

compete with them, but we can compete *cooperatively* with them in virtue of doing so in the inner layers of our agency. It is easy to see how this pattern fits many games.

As Ferrero and Kukla point out in their replies to Nguyen, actions in real life outside the context of games often seems to be structured similarly. It is extremely rare that acting on any of our outer layer-aims does not take place under important constraints, whether material, cultural, or interpersonal. I am trying to write, but I get hungry (material constraint), so I decide to make a sandwich rather than eat bread without toppings (cultural constraint), but to get myself a sandwich I need to buy the bread and toppings from the grocery store (interpersonal constraint).

We may then map List & Pettit's terminology onto Nguyen's. Individual agents may have their outer-layer aims, but to realize them, they may very well decide to take up the constraints implicit in corporate agency. This is what agent-switching involves. Presumably this can sometimes happen from an outside perspective, in the sense that an agent actively compares the aims of a corporation with their own and tries to judge whether they are instrumentally helpful. But on List & Pettit's picture, what primarily happens is that individual agents identify with the corporate agent and then engage immersively with its aims after taking up its point of view and mental state tokens. These may override or even stand to the detriment of individual agents' own long-term aims, much like how someone playing a game sometimes may engage with its aims to an extent which overrides or is detrimental to their own long-term aims. Nevertheless, agent-switching is not *prima facie* odd. It is just a feature of our ordinary lives.

(2) The Argument

I shall, however, argue that List & Pettit-style agent-switching cannot occur. Here is my argument:

- (1) Individual agency is plight inescapable.
 - (2) If individual agency is plight inescapable, individual agents are continuously faced with fully exercising their own agency (absent external limits at the time of its exercise).
 - (3) If individual agents are continuously faced with fully exercising their own agency (absent external limits at the time of its exercise), individual agents cannot switch into participating in other forms of agency with irreducible mental states.
-
- (C) Individual agents cannot switch into participating in other form with irreducible mental states.

The argument is valid. If it is sound, it counts against agent-switching and realism about corporate agency. I shall defend the premises of the argument in one section each below. Then I shall wrap up by considering the implications for theories of corporate agency.

(3) Premise (1) – Individual Agency is Inescapable

Premise (1) says that individual agency is plight inescapable. What does that mean? It is inspired by Korsgaard, so let us start with her account (2009, pp. 1-2). Her characterization of plight inescapability is very brief, but the core point is that agents continuously are faced with performing new actions: even choosing not to act is still to act, as that, too, involves acting. Hence action is inescapably our plight. In other words, it is *plight inescapable*.

This point plays a key theoretical role in motivating Korsgaard's grander Kantian constitutivist theoretical project (Ferrero, 2019). For Korsgaard, the constitutive features of action and agency are *inter alia* the categorical and hypothetical imperatives, and our constitutions therefore explain why we are committed to morality and rationality (Korsgaard, 1996; 2009). But these norms hold for us independently of what we would want to do, feel like doing, or desire. We cannot shirk on them just by not wanting, feeling, or desiring to follow them.

To explain why we cannot shirk, a key background assumption is that they apply to all exercises of agency, including both intentional actions and intentional omissions.⁶ This is where plight inescapability comes in. If we would not continuously face exercising our agency via the imperatives, we would not be subject to them on all relevant occasions. But we are continuously faced with acting on them, so we are continuously subject to them. So plight inescapability is part of the explanation of why the constitutive features of agency have force over us independently of what we want, feel, or desire.

However, plight inescapability is not just a theoretical background assumption for Korsgaard. The reason it slots very well into her picture is that she is right, at least in broad outline: plight inescapability is a phenomenological datum. This is so because it is true that agents are continuously faced with exercising agency (at least absent external limits – I shall return to this point in section (4) below). Say you fail to do what you are supposed to do. Then you are faced with doing something else – better – in literally the next moment in which you can act. Or say you get knocked unconscious, but you wake up, and no lasting damage has been done to you. Then you are back to acting at full capacity again. So exercising agency seems plight inescapable.

It is important to emphasize the independence of this point from several other things, however. First, Korsgaard's phenomenological insight into the conditions under which we exercise agency should not be thought of as equivalent to the phenomenological experiences that each agent has

⁶ Talking in terms of 'exercises of agency' implicitly solves a potential worry. One might think that intentional omissions are not actions, but surely one can intentionally omit acting (Clarke, 2014)? Ok, but doing that still involves exercising agency. I return to the locution of 'exercising agency' in section (4) below.

each time they act. It is of course true that we often fail to act, or act half-heartedly, but that does not mean that we are not continuously faced with acting at full capacity again afterwards.

Second, whether agency is plight inescapable is orthogonal to the question of which normative commitments agency might involve, whether constitutively or relative to an agent's capacities. Therefore, it is not restricted to Korsgaard's theoretical framework, or even to Kantianism more generally. It is quite conceivable that agency involves something else than the Kantian imperatives, but it is still inescapable in the sense that whatever it features is something that an agent is continuously faced with. This means that an agent is inescapably committed to *whatever* normative commitments she has. It may even be true that corporate agents are faced with exercising whatever it is that some particular form of corporate agency may involve, but that does not matter from an individual member's perspective: they are faced with exercising *their own* agency.⁷

Third, the phenomenological point is independent of other types of inescapability. Different constitutivists have emphasized different kinds of inescapability in different contexts. Some constitutivists have emphasized them to explain normative force: for example, *psychological* inescapability, which means that there is a motive which always affects an agent (Velleman, 2009) or *standpoint* inescapability, according to which there is no standpoint outside agency from which one may ask normative questions (Ferrero, 2019). Others have discussed it in the context of the shmagency objection, according to which it might seem pre-theoretically possible to escape the commitments of agency by being some other type of entity than an agent: in other words, a 'shmagent' (Enoch, 2006; 2011; Leffler, 2019). Here, the most common response has been to argue that agency is *dialectically* inescapable, viz. that trying to escape agency is self-defeating in virtue of the fact that all actions (and intentional omissions) involve agency and that even reflecting on escaping it, let alone trying to do so, still is to exercise it (Ferrero, 2009; 2018; cf. Leffler, 2019 for criticism). It may very well be true that action or agency are escapable or inescapable in some other sense or senses than plight inescapability, but those are not the types of inescapability currently under discussion.

In sum, plight inescapability is a phenomenological datum but need not always be experienced by all agents, and it is neither necessarily connected to Kantianism nor the same thing as other types of inescapability. Yet to see its philosophical implications for agent-switching, we will want a full

⁷ Korsgaard sometimes writes as if Kant's conception of agency and Plato's in *The Republic* are identical, so individuals and collectives face organising themselves in the same way (Korsgaard, 2009, ch. 6-7; 2014). So shouldn't we treat them analogously? I argue against that below, but here I want to emphasize that this does not matter from the individual agent's perspective.

account of it, rather than just the outlines above. I now turn to provide one, *ipso facto* defending premise (2).

(4) Premise (2) – Plight Inescapability as Continuously Exercising Agency Fully

Premise (2) extends plight inescapability. To recapitulate, it reads: '(2) If individual agency is plight inescapable, individual agents are continuously faced with fully exercising their own agency (absent external limits at the time of its exercise).' How so?

There is not much systematic work on how plight inescapability should be understood. While Korsgaard has got *something* right, it is unclear what exactly she has got right. And she does not characterize plight inescapability in sufficient depth to give us a full view. For the most theoretical characterization of plight inescapability in the literature, we may instead turn to Luca Ferrero's account.

For Ferrero, plight inescapability consists of exercising a two-way power. Powers have recently often been invoked to interpret agential abilities, often even with appeals to a two-way power to act or refrain from acting; indeed, that is exactly how Steward (2012) uses a two-way power in a defence of free-will libertarianism. Similarly, Ferrero takes the power involved in plight inescapability to be the 'power to either make an antagonistic intervention in the natural course of events or to refrain from such intervention' – unless the agent is in 'a dormant state' (Ferrero, 2018, p. 135).

This view improves on Korsgaard's brief categorization of plight inescapability. It is formulated in more familiar metaphysical language, it allows us to understand what an agent is doing when she is not acting (she is refraining or in a dormant state), and it allows for cases where an agent is not exercising her power at all by being in a dormant state: etymologically, a 'sleeping' state. Ferrero's suggestion is, however, somewhat impressionistic. This opens the door for two objections.

First, it seems better to describe the plight inescapability of agency without the theoretically contentious notion of a two-way power. This is so because, quite generally, the metaphysical connection between agency and various types of potentialities, whether they are powers, capacities, abilities, dispositions, or something else, is rather tricky to make sense of. A less theoretically assuming way to describe what happens here is to say that agents may exercise their own agency. This allows the reader to fill in with her preferred metaphysics. And it allows us to, for present purposes, treat intentional omissions like intentional actions, as per footnote 6 above.

Second, the notion of a 'dormant state' is too reductive. I am happy to accept that a sleeping agent ordinarily does not exercise their two-way power, but Ferrero does not explicitly characterize

dormant states in more depth. However, agents may fail to exercise their agency in many ways, not just when they are asleep: for example, a rock could drop in someone's head, they could have tied themselves to the mast, or have failed to develop skills that are necessary for some action they are trying to perform. We need a more fine-grained explanation of them than 'dormant states.'

To avoid these concerns, plight inescapability should rather be thought of in terms of agents continuously facing the full exercise of their own agency (absent external limits at the time of its exercise). Even more explicitly: as long as agents remain agents, they are always in a position to fully exercise their own agency, either in virtue of acting or intentionally omitting action, absent some limit external to the exercise of agency at the particular time of its exercise.

Here, a 'limit' is whatever serves to block the exercise of agency. It can be entirely external to the exercise of agency, such as a blow to the head that knocks one unconscious. Alternatively, it could have been brought about either deliberately or non-deliberately by the agent, such as by going to sleep at t_0 , therefore limiting her from exercising her agency at t_1 , or by not developing her skills or abilities adequately, limiting her from acting in the way she is attempting. This view avoids the problems for Ferrero. It is not formulated in terms of potentiality concepts, and it is more informative than characterizations of non-agency in terms of a dormant state. Hence, my account improves on his.

However, another feature of my account needs more emphasis, for I shall put it to good use below. What is plight inescapable is not just being faced with continuously exercising agency absent external limits at the time of the exercise, but *fully* exercising it. Exercising agency fully means exercising agency fully successfully: one can either φ or fail to φ relative to the standards of success of some particular exercise of agency, which in turn determine what *the thing to do* is at some particular time. Hence, what one faces when one faces exercising agency fully is to do so successfully with respect to the standards of success that determine what the thing to do is.

What does that mean? 'The thing to do' outlines the correct answer to the question of 'what to do' raised in deliberation (Southwood, 2018; cf. Gibbard, 2003). Deliberation concerns settling what to do, at any time, where that is just one of the set of many possible exercises of agency available to us. One might then say that deliberation raises the question of what to do, and the thing to do is the correct answer about what to do. This is so whether we actively deliberate or not.

The standards of success for our actions are set by our normative commitments: for example, the standards constitutive of agency, or maybe our reasons. In the light of these, we need to determine how to exercise our agency at each point in time, for we are faced with settling what the thing to

do is in the light of the standards of success that determine that. Our normative commitments may, however, often come into conflict. But as we need to settle on the thing to do, we are also faced with balancing them out to the extent possible: for example, by acting on the weightiest reasons, or one of several possible sufficient reasons.⁸ These are likely to be the things to do. Accordingly, what we are faced with is exercising agency fully, where that includes exercising it successfully in the light of the normative commitments that determine what the thing to do is (absent external limits at the time of its exercise), and the thing to do just is the correct answer to the question of what to do in the light of our varying commitments.

Even so, it is possible to exercise agency without doing so fully. Perhaps something else than one's agency gets in one's way. Again, maybe a rock in the head leaves one unconscious. Or, more theoretically interestingly, the pain the rock causes could limit one from performing some action under the circumstances under which one acts. With a headache, I might maybe go on a walk instead of finishing writing this paper, even though that is my longer-term intention. Or maybe some agent has constrained her agency at t_1 with her agency from t_0 – perhaps, again, in virtue of deliberately falling asleep. But in all these cases, there is something external to the agency of the agent *at the time of the exercise of agency* which is limiting her from exercising her agency fully (at that time).

Hence, I think plight inescapability should be understood in terms of an agent's being continuously faced with fully exercising their own agency (absent external limits at the time of its exercise), where that is characterized as I just have done. If this view is on the right track, premise (2) is justified. For it says that if agency is plight inescapable, agents face fully exercising their own agency (absent external limits at the time of its exercise). And I have defended an account of plight inescapability which involves exactly this.

To clarify and justify my account further, I shall conclude this section by dealing with some potential objections. They are familiar from the literature on constitutivism but are now adapted to help explicate plight inescapability further. A first worry is that agency still might seem escapable. Plight inescapability does not show that it is impossible to stop being an agent. And, yes, it is quite possible to escape agency, in some ways. But plight inescapability is a very specific kind of inescapability, as I briefly hinted at in section (3). It is independent of other types of inescapability – and of many ways of escaping agency: it is logically, metaphysically, and nomologically possible to stop being an agent, for example by dying. But plight inescapability is a *practical* kind of

⁸ Of course, how we are to do that can often be a very complex question. Fortunately, we need not determine how to settle on exercises of agency in the light of conflicting normative commitments in this paper – but we *do* need to settle them in practice.

inescapability, for one continuously faces exercising agency by being in a position to do so again as long as one is an agent. I take this to be Korsgaard's original phenomenological insight, as I outlined above when defending premise (1) in section (3): absent external limits, agents face exercising their agency fully. This is so in virtue of their facing the possibility of doing so. In fact, even so-called shmagents – or non-agents who at best approximate having the constitutive features of agency, without having them – also face exercising whichever non-agential ('shmagential') features they have, too. For they face the same situation of working out what the thing to do is. So plight inescapability is not a feature of agency per se: it is a feature of the conditions under which agency (or shmagency) is exercised. This is something we always face, in practice, whether or not we are agents or something like them.

Another worry can be distilled from Rovane (1997, ch. 4; 6). Against the idea that it seems like individual humans appear to have a rational point of view that they cannot shake off and which therefore is inescapable, she argues that it is important to notice how different rational points of view – which for her are those of persons, who can be individual or corporate – come into being over time. We are not born with rational points of view, and we can create them. It might therefore seem like it is contingent whether we inhabit some particular rational point of view. We can create them at will.

This worry is, however, irrelevant for plight inescapability. The plight inescapability of agency is faced by all agents in the light of the normative commitments that they *already* have. We may create any agents we like, by whichever means, but normative commitments are still plight inescapable for those who have them. This is because everyone continuously faces fully exercising them.

But could one just not do something less than exercising agency *fully*? Perhaps one could *try* to do something rather than do it. However, the point about facing exercising agency *successfully* above implicitly responds to this worry. Trying to exercise agency is not, by itself, to do so fully. But a successful trying *is* a successful exercise of agency. If I successfully try to pour myself a glass of wine, then I *ipso facto* pour myself a glass of wine. If I successfully try not to pour one as I shouldn't have any more, I *ipso facto* omit. Hence, trying is another way of exercising agency, and one is faced with exercising it fully.

This point also contributes to responding to a related worry. For one might think that one could exercise agency by *failing* to act, and hence not exercise it at all. This is the famous problem of bad action for constitutivists (cf. Clarke, 2000; Lavin, 2004). For example, if agency constitutively involves acting on norms – the categorical and hypothetical imperatives, say – it should be possible to fail to live up to both but still act. Suppose it is categorically wrong to lie about where one was

last night, and I intend to lie about it by saying I was at home rather than at the pub, but I lie about it by saying that I was at the bar instead, I seem to both fail to live up to the norm that it is wrong to lie and that of taking means to my ends. But I still act, even though I fail to live up to both norms.

There are, however, familiar responses to this too. Some think agency comes in degrees and therefore can be more or less enacted. This is Korsgaard's (2009) view: acting involves abiding by the constitutive standards of agency to at least some minimum threshold, but once one is over it, agency can be exercised more or less (and more or less *well*). I prefer a twist on that idea: it involves thinking that action comes in different *kinds*, so an agent can perform different kinds of actions depending on how they exercise their agency.⁹ But the details are unimportant here. What matters is that a defender of plight inescapability can avail herself to various solutions.

(5) Premise (3) – The Tension

I have now characterized plight inescapability and addressed some potential concerns. It is time for another reminder. Premise (3) reads: 'If individual agents are continuously faced with fully exercising their own agency (absent external limits at the time of its exercise), they cannot switch into participating in other forms of agency with irreducible mental states.' Come again?

This premise can be defended because the commitments of individual agency are plight inescapable for individual agents. We can spot the tension between individual and corporate commitments most clearly when we consider how corporate collectives often make commitments that members do not endorse. As per List & Pettit's account of agent-switching from section (1) above, it is quite common that an individual who enacts the will of a corporation may end up acting against their own aims or desires. To make the price-raising example from above more vivid, we may think that an employee in a grocery store may want to lower the price of baby formula to make it more accessible to poor families during a shortage, whereas the owner might decide to raise prices to profit from the shortage.

In Nguyen's terminology, the employee may have an outer layer-aim of helping the poor, but in virtue of being an employee of the corporation, they face the constraint of acting in line with the corporate decision in the inner layer of their agency. But this is not just because they are anticipating the corporation's aims as something external to their own. Rather, because the employee

⁹ See [author's work].

presumably will enact the corporation's aims when they are acting in their professional role and meeting customers at work, they will be taking up and enacting the desire of the corporation.

However, individuals cannot fully exercise their own agency if they enact the aims or desires of the corporations if these are construed as irreducible mental states over and above their own. In particular, if the corporate commitments also are different from those of the individual employee, and the individual agent would attempt to enact the decisions of the corporate agent, the employee would not adhere to the normative commitments of their own agency. But the employee is, *qua* individual agent, faced with fully exercising them. This means that the employee cannot switch while maintaining their inescapable commitments: they are unable to enact the policy they disagree with in case they are facing up fully to the commitments of their own agency. And that is plight inescapable.

In fact, the employee would not be able to switch to and enact the corporate agent's desires even if they would happen to align and support the same course of action as their own. Assume, contrary to the above, that both the employee and the corporation desire to raise prices. Even then, the individual agent is unable to enact the aims of the corporation by identifying with and implementing its desire token. This is because what is plight inescapable for an agent is to exercise their own agency, including its normative commitments, rather than those of another agent. When there is an alignment in goals between different agents, these still rather consist of metaphysically distinct mental state tokens on part of different agents with similar content.

The relations between different mental states here can be given different – albeit sometimes complementary – explanations on different metaphysical accounts of action. For example, a typical constitutivist way to get at the potential tension here is Kantian. If it is constitutive of agency to follow the categorical imperative, as Korsgaard (1996; 2009) thinks, a maxim that permits raising the price in a shortage may very well not be universalizable.¹⁰

Leaving ambitious constitutivist territory behind, on most accounts agency, standard agents have the capacity to respond to reasons (Smith, 1994; Dancy, 2000) or to enact longer-term plans (Bratman, 1999). If the individual employee has strong reasons or plans that indicate that prices ought to be lowered, then these are the sorts of things that would stand in tension with taking up the corporate perspective.

¹⁰ Note that this is just an example view. I will not try to argue that one cannot be: making good sense of maxims and universalization would be far beyond the scope of this paper.

Yet these types of views also make it easy to see how goals can align between individuals and corporations: it is quite possible that the individual employee could have reasons or plans to increase prices. People participate in collective actions for all kinds of reasons: perhaps a living wage is very important for them as they need it to feed their family, even though they in principle agree that prices should not be raised. Such an agent will have their aims only grudgingly aligned with the corporation. They will then have to treat its goals as constraints on the inner layers of their agency, even though their outer-layer aims point in another direction. Alternatively, the employee could be extremely self-regarding: maybe increasing prices will raise profits, which could earn them a bonus, which they care more about than the well-being of others.¹¹ Here the alignment might not seem constraining to them from the inside, but there would still be *potential* constraints on how to exercise their aims set by what they need to do to act in their role in the corporation. Quite independently of what one's metaphysics of agency might be, there may therefore be tensions or alignments between acting on one's own agency and the corporate agency one is part of.

Nevertheless, *if* enacting a normative commitment to lower prices is important for the employee, then they would no longer exercise their own agency if they were to take up the perspective of the corporation. But they cannot do that because their own agency is inescapable. Nor can the employee do that if she were to desire the same thing as the corporation, as agents cannot switch to acting on irreducible corporate mental state tokens they do not have themselves: their agency is inescapable in that case too. There is always a tension between individual attitudes and corporate ones, as the former are plight inescapable. So premise (3) is justified.

A possible worry here is, however, that this apparent tension may assume too little on part of individual agents. For example, it might look like individual agents cannot, in any sense, identify with corporate agents. This may seem odd, as we clearly sometimes do act closely together: social interaction does not always consist of individualistic game-theoretical strategical maximizing. But I am not arguing against that. Rather, the worry is that the normative commitments of an individual's agency (at the time of its exercise) are plight inescapable for that agent, which shows that the agent cannot switch into possessing other mental state tokens than their own (such as corporate ones). *That* type of identification is ruled out. But there is no tension between having an aim and sharing it with others as long as it is still one's mental state token. One may even identify with a corporate agent in other ways: for example, by believing one is part of one, or by desiring that it does well.

¹¹ I assume here that self-regarding desires indeed may ground reasons. Philosophers who think who moralize reasons more, whether ambitious constitutivists like Korsgaard or externalists like Dancy, may disagree. But, for the purpose of this example, it is at least intelligible how a bonus may provide *some* reason.

But are our powers of acting not greater still? For example, we are able to decide to treat some beliefs or reasons hypothetically, *as if* they are what we believe or what we have reason to do, whether or not we think they are. Or we can pretend: we can act *as if* we were members of a corporate entity, whether or not we are.¹² Why can we then not treat corporate attitudes *as if* they are our attitudes, or pretend that they are? Well, we can. But it is still an individual agent who does that, in the sense that hypothetical attitudes or pretending still are things the individual agent has or does. They do not entail switching into *actually* having others' mental state tokens.

Another potential worry: is participation in corporate action the kind of limit for the exercise of individual agency that would indicate that the antecedent of premise (3) will not lead to the consequent? That would take us close to the position of Rovane (2014). She thinks that agents can participate in either individual or collective agency, even though taking up a collective perspective *diminishes* the perspective of the individual agent. If their perspective gets in that way diminished, maybe they can take up a conflicting perspective?

However, actions in corporate settings do not normally become diminished. When we act together with others, even when we disagree with a joint decision, the action is nevertheless ordinarily considered fully our own.¹³ Think again of the case of immersion in games. If an ice hockey player does not want to score a goal because doing so would relegate the opposing team from the league and the opposing team was their childhood favourite, but the player's current team aims to win, the player makes an effort to score a goal in the game, and they do score it, it seems odd to say that their agency when scoring would be diminished. Rather, the player is likely to be acting under the constraints of the game which feature in the inner layer of their agency, and quite possibly even be immersed in them. But that is still fully their own agency. They are acting on their own aim – to win with their team – even though they to some extent may regret it if it is successfully enacted. But that is not unusual. If we tie ourselves to the mast, we may feel some regret about being tied to it, whether we are acting in a corporate context or not.

To summarize: if individual agents are continuously faced with fully exercising their own agency (absent external limits at the time of its exercise), they cannot switch into participating in other

¹² Thanks to {X and Y} for discussion of these issues.

¹³ As an anonymous reviewer helpfully has emphasized, there are also cases where someone's agency in a group context *is* diminished: for example, if the group is a cult, or they are subject to propaganda. But individual agents can nevertheless not switch into possessing the mental state of a larger group; plight inescapability makes that impossible, since it makes it impossible to switch whether or not the agents disagree. Without being an expert on either, I rather suspect cults or propaganda are likely to make an agent's agency *align* with the goals of the group, presumably by changing the content of their mental states.

forms of agency with irreducible mental states, in the sense that they have distinct mental state tokens. That is premise (3). So premise (3) is defended. Hence (C) follows.

(6) The Problem

We have learnt that individuals cannot switch into other forms of agency than their own. This is because their own agency is plight inescapable. But this result does not just say something about the psychology of agent-switching. It also generates a problem for realism about group agency, where the mental states of the group irreducibly stand over and above those of their members. This problem suggests that we probably ought to prefer some reductive views of corporate agency to realism – though which view might be best is an open question: we might perhaps think of corporate agents in terms of shared intentions (Bratman, 2013; 2022), following constitutive rules (Pauer-Studer, 2014), having shared ends (Miller, 2003), groups acting for motivating reasons without group minds (Silver, 2022), or something else.¹⁴

The problem here is a problem here for all kinds of realists who think corporate mental states stand over and above those of their members. This is true on the received realist view introduced in section 1 above (Björnsson & Hess, 2018; French, 1979; Hess, 2020; List & Pettit, 2011; Rovane, 1997, 2014; Tollefsen, 2015). As mentioned, realism comes in different varieties, but they all take the attitudes of a corporate agent to depend on the attitudes of individual participants take group attitudes and actions to differ from individual ones. This means that the view allows for tensions between individual and corporate attitudes regarding some aim.

Views from the literature on joint intentions also have similar issues (Gilbert, 2009; Tuomela, 2013). Indeed, Gilbert and Tuomela count as realists in my sense, though their views differ slightly from the most common ones. For Gilbert, there are joint intentions that can differ from those of individuals (Gilbert, 2009): individuals constitute plural subjects in virtue of their joint commitments, which generates mutual obligations for the participants. The latter need not be true on standard realist views, but what matters here is that the joint subjects can have mental states – including intentions – that differ from those of the committed individuals. For Tuomela, a group agent is a causal-functional entity which can cause events based on desires and intentions that differ from what the individuals who constitute them desire or intend for themselves (in the ‘I-mode’ rather than the ‘we-mode’). Even though Tuomela does not think that group agents have mental states in a full sense, as they lack phenomenal consciousness, we can still attribute mental states to them in a functional sense. Hence, in both Gilbert’s and Tuomela’s frameworks, corporate agents

¹⁴ The argument in this paper is even compatible with eliminativism about group agency, but it does not entail it. It is an argument against realism, not against group agency.

can have aims (for example, intentions) that differ from the aims of their individual members. Accordingly, my worry appears for them too.

But what exactly is the worry? Exercising one's own individual agency fully is plight inescapable (absent external limits at the time of its exercise) *even* when one is acting together with others. This ought to be explained. But views that detach the aims, whether desires or intentions, of individual agents from corporate agents entail that individual agents are no longer acting on their own inescapable commitments when they act on attitudes held by the corporate agents they are members of.

Reductive views do not have this problem. Here, individuals' mental states may be identical to those of a corporate agent – or not even play a key role in the explanation of corporate action at all. On these views, an individual agent need not switch into having other mental state tokens than their own when acting in a corporate context. As such, they will be able to act on their own inescapable commitments. Hence, a view along these lines seems preferable to realism.

There is one final problem. Above, I briefly addressed a worry about the link between the antecedent and consequent of premise (3): namely, that individual and corporate agencies seem to stand in tension and can end up diminishing each other (Rovane, 2014). The reply was that we ordinarily are inclined to think of the actions of an individual agent (in my case, an ice hockey player) in a corporate setting as their own action even though they might voice some regret about it. This point is analogous to how a long-term intention on part of an individual agent can stand in tension with capricious and momentaneous desires.

But Rovane's view suggests a deeper worry. What happens when an agent seems to intentionally and deliberately take corporate and individual perspectives to differ? Perhaps an agent at t_0 expresses an attitude which they take to belong to a corporate agent, but then at t_1 expresses a different attitude which they claim to be their own. Returning to the baby formula example, perhaps the owner of the grocery store chain might claim at t_0 that the corporation wants to raise prices, but at t_1 that that is not the owner's own attitude. The owner might claim to prefer that prices should not be altered or even decreased. If we should take such perspective shifts at face value, taking agent-switching to be a problem seems to leave our theories of corporate agency impoverished.

But individual agents can be wrong. We can plausibly make sense of the owner's actions in terms of his or her own mental states. I shall outline some possibilities for how that might work: depending on how we spell out the details of the case, they could all be true. We can, then, rather

easily give reductive explanations to accommodate apparent perspective shifts once we attend to the facts of the matter, whichever they happen to be.

Two possibilities here involve attempts to save face. Perhaps the owner changes their mind between t_0 and t_1 , and then at t_1 tries to appeal to the attitudes of the corporation as an excuse for having said something awkward when suggesting that prices ought to be raised at t_0 . An even less conscientious owner could simply lie. Perhaps the owner continuously believed that prices ought to be raised but then, at t_1 , realized their t_0 awkwardness when stating their view explicitly at t_0 , and so lies about their view at t_1 .

The owner's judgement shifts may also be explained with some more complex mistake. *Akasia* is one possibility. Perhaps they are a pushover egoist who thinks it right to raise the price at t_0 , but then gets overwhelmed by a desire to appease whomever they are talking at t_1 . Their desire to appease could then overtake the egoism. Alternatively, the owner might be a pushover altruist who thinks the right thing to do even at t_0 is to lower prices, but a desire to appease others – such as corporate board members – might overwhelm them at t_0 , leading them to raise prices and then explain the raise in terms of the corporate will to someone else at t_1 .

Finally, perhaps there could be a conflict that arises because the owner plays the corporate game too hard. Possibly, immersed in telling a kind of profit-maximizing quasi-fiction at t_0 which they are committed in virtue of having a certain corporate position, even though it does not represent a genuine attitude, they may have 'bluffed' at t_0 , analogously to how bluffing in poker is (reasonably) fair game. Their mistake would then be to speak to someone else – either inside the corporation or not – while staying too far in game without taking moral considerations into account.

As such, there are many potential explanations for why an individual might say they are switching between standpoints. They vary with how we cash out their psychology. Hence, the fact that individuals sometimes may refer to taking up different standpoints does not establish that they in fact do. There are many other, often deeper, competing psychological explanations of what happens when an agent seems to take up a corporate perspective that differs from their own. So views according to which individual and corporate agency can come apart in terms of the mental states they feature still have a problem: individual agents continuously face exercising their own agency fully even when acting in corporate contexts.

(7) Conclusion

In this paper, I have argued that plight inescapability creates a problem for accounts of corporate agency where these have aims that consist of irreducible mental state tokens. This is because plight

inescapability rules out the kind of agent-switching that such accounts seemingly are committed to. To argue this, I outlined the phenomenon of agent-switching in section (1), the problem with agent-switching in section (2), defended its premises in sections (3)-(5), and finally outlined the problem for defenders of group agent realism in section (6).

All this generates a metaphysical challenge for group agent realists: is it plausible to assign groups mental states that individual agents are supposed to be able to switch to when they are individually inescapably committed to their own mental states? More will have to be said here – with two *caveats*. First, the problem does not say anything about which more reductive approach to corporate agency we might want to prefer. Second, even though much work on social ontology is motivated by normative considerations about things like group responsibility, group epistemic norms, or politically salient actors, the criticism leveraged at group agent realism is descriptive rather than normative. Hence, integrating its message with normative concerns will be an important theoretical task. It will, however, be a task for later and for elsewhere.

References

[Author's work]

Björnsson, G. & Hess, K. 2017. Corporate Crocodile Tears? On the Reactive Attitudes of Corporate Agents. *Philosophy & Phenomenological Research*. 94(2), pp. 273-298.

Bratman, M.E. 1999. *Intention, Plans, and Practical Reasons*. Stanford, CA: CSLI Publications.

--- 2013. *Shared Agency: A Planning Theory of Acting Together*. Oxford, UK: Oxford University Press.

--- 2022. *Shared and Institutional Agency*. Oxford, UK: Oxford University Press.

Brown, J. 2022. Group Belief for a Reason. *Aristotelian Society Supplementary Volume*. 96(1), pp. 1-22.

Clark, P. 2001. Velleman's Autonomism. *Ethics*. 111(3), pp. 580-593.

Clarke, R. 2014. *Omissions: Agency, Metaphysics, and Responsibility*. Oxford, UK: Oxford University Press.

Dancy, J. 2000. *Practical Reality*. Oxford, UK: Oxford University Press.

Enoch, D. 2006. Agency, Shmagency: Why Normativity Won't Come from What is Constitutive of Action. *Philosophical Review*. 115(2), pp. 169-198.

--- 2011. Shmagency Revisited. In: Brady, M. ed. *New Waves in Metaethics*. New York, NY: Palgrave MacMillan, pp. 208-233.

Ferrero, L. 2009. Constitutivism and the Shmagency Challenge. In: Shafer-Landau, R. ed. *Oxford Studies in Metaethics*. 4, pp. 303-332.

--- 2018. Inescapability Revisited. *Manuscripto*. 41(4), pp. 113-158.

--- 2019. The Simple Constitutivist Move. *Philosophical Explorations*. 22(2), pp. 146-162.

- 2021. Games and the Fluidity of Layered Agency. *Journal of the Philosophy of Sport*. 48(3), pp. 344-355.
- French, P.A. 1979. The Corporation as a Moral Person. *American Philosophical Quarterly*. 16(3), 207-215.
- Gibbard, A. 2003. *Thinking How to Live*. Cambridge, MA: Harvard University Press.
- Gilbert, M.P. 2009. Shared Intention and Personal Intentions. *Philosophical Studies*. 144(1), pp. 167-187.
- Hess, K. 2020. Assembling the Elephant: Attending to the Metaphysics of Corporate Agents. In: Bazargan-Forward, S. & Tollefsen, D. eds. *The Routledge Handbook of Collective Responsibility*. New York, NY: Routledge. pp. 113-126.
- Korsgaard, C.M. 1996. *The Sources of Normativity*. Cambridge, UK: Cambridge University Press.
- 2009. *Self-Constitution: Agency, Identity, Integrity*. Oxford, UK: Oxford University Press.
- 2014. The Normative Constitution of Agency. In: Vargas, M. and Yaffe, G. eds. *Rational and Social Agency: The Philosophy of Michael Bratman*. Oxford, UK: Oxford University Press, pp. 190-214.
- Kukla, Q.R. 2021. Sculpted Agency and the Messiness of the Landscape. *Analysis* 81(2), pp. 296-306.
- Pauer-Studer, H. 2014. A Constitutive Account of Group Agency. *Erkenntnis*. 79(S9), pp. 1623-1639.
- Pettit, P. 2023. *The State*. Princeton, NJ: Princeton University Press.
- Lavin, D. 2004. Practical Reason and the Possibility of Error. *Ethics*. 114(3), pp. 424-457.
- Leffler, O. 2019. New Shmagency Worries. *Journal of Ethics and Social Philosophy*. 15(2), pp. 121-145.
- List, C. & Pettit, P. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford, UK: OUP.
- Miller, S. 2003. *Social Action: A Teleological Account*. Cambridge, UK: Cambridge University Press.
- Mitova, V. 2022. The Collective Epistemic Reasons of Social-Identity Groups. *Asian Journal of Philosophy*. 1(2), pp. 1-20.
- Nguyen, C.T. 2020. *Games: Agency as Art*. Oxford, UK: Oxford University Press.
- Rovane, C.A. 1997. *The Bounds of Agency: An Essay in Revisionary Metaphysics*. Princeton, NJ: Princeton University Press.
- 2014. Group Agency and Individualism. *Erkenntnis*. 79(S9), pp. 1663-1684.
- Silver, K. 2022. Group Action Without Group Minds. *Philosophy & Phenomenological Research*. 104(2), pp. 321-342.
- Smith, M. 1994. *The Moral Problem*. Oxford, UK: Blackwell.
- Southwood, N. 2018. Constructivism and the Normativity of Practical Reason. In: Jones, K. and Schroeter, F. eds. *The Many Moral Rationalisms*. Oxford, UK: Oxford University Press, pp. 91-109.
- Steward, H. 2012. *A Metaphysics for Freedom*. Oxford, UK: Oxford University Press.

NB. Author's accepted copy. The paper is available, with editorial changes, at: <https://doi.org/10.1111/phib.12317>.

Strohmaier, D. 2020. Two Theories of Group Agency. *Philosophical Studies*. 177(7), pp. 1901-1918.

Tollefsen, D.P. 2015. *Groups as Agents*. Hoboken, NJ: Wiley.

Tuomela, R. 2013. *Social Ontology: Collective Intentionality and Group Agents*. Oxford, UK: Oxford University Press.

Velleman, J.D. 2009. *How We Get Along*. Oxford, UK: Oxford University Press.