



# What kinds of groups are group agents?

Jimmy Lewis-Martin<sup>1</sup> 

Received: 18 June 2021 / Accepted: 1 June 2022 / Published online: 30 June 2022  
© The Author(s) 2022

## Abstract

For a group to be an agent, it must be individuated from its environment and other systems. It must, in other words, be an individual. Despite the central importance of individuality for understanding group agency, the concept has been significantly overlooked. I propose to fill this gap in our understanding of group individuality by arguing that agents are autonomous as it is commonly understood in the enactive literature. According to this autonomous individuation account, an autonomous system is one wherein the constituent processes of the system actively produce and sustain that self-same system, which will run down or fail if any of these constituent processes cease. This definition of autonomy provides us with a precise and operational account of the individuality of group agents. I will then compare this account to those of Carol Rovane and Raimo Tuomela to argue that it offers the best explanation of what kinds of groups are group agents.

**Keywords** Autonomy · Enactivism · Group agency · Individuality · Individuation · Social ontology

## 1 Introduction

What kinds of groups are group agents? Despite the recent upsurge of interest in the nature of group agency (List & Pettit, 2011; Pauer-Studer, 2014; Rovane, 2019; Tollefsen, 2002, 2015; Tuomela, 2013), the issue of determining what kind of group is the right kind remains a contested matter. To address this question, I will focus on

---

✉ Jimmy Lewis-Martin  
jimmy@uow.edu.au

<sup>1</sup> University of Wollongong, Wollongong, NSW, Australia

establishing what it is that makes a group an *individual* capable of agential activity.<sup>1</sup> Xabier Barandiaran, Ezequiel Di Paolo, and Marieke Rohde (2009) term this the individuality criterion of agency: for a system to be an agent, it must be an individual, distinct from its environment and other systems (p. 369).<sup>2</sup> To put it differently, for a group to be an agent, there must be some internal coherency to the system in order for it to be considered an agent, just as we would expect of any agent.

If the group is itself the system to which its actions are attributable, it cannot be reducible to its parts, nor can it merely be the sum of its parts. A group that is reducible to its parts is a group in a metaphorical sense only. In these cases, such as when we talk of what ‘the market’ demands or what ‘the people’ want, we are making use of metaphorical shorthand to summarise a pattern in personal beliefs and behaviours (List & Pettit, 2011, pp. 2–3). Examples of groups that are nothing other than the sum of their parts are reading groups, friend groups, and romantic couples. In these cases, the key feature is shared intentionality (see Bratman 1992; Gilbert, 2009). When we make plans to go to dinner with friends, we form a shared intention. The point is not merely to go to dinner – it is to go to dinner *together*. Likewise, the point of a reading group is not just to read a book, but to read a book *together*. These groups exist when people interact with each other in distinct ways, namely when they share intentions. But for these groups, the actions of the group are reliant on the attitudes and behaviours of the group members.

What is distinct about a non-metaphorical group *agent* is that the group, rather than its members, is the source of activity. It sets the agenda, regardless of the personal views of the singular agents who enable it. As Tuomela (2013) claims, group agents are ontologically real insofar as they causally influence action in a top-down manner (p. 5; p. 47). Groups depend on their members because their members are, in many cases, the material that makes up much of the group. However, those members are constrained in their potential actions by the group. Were they not, then the group itself would not be the agent and ‘group agency’ could only ever be a metaphorical concept. It is therefore important to understand group agents on their own terms, i.e., directly, rather than in terms of their members.

The attitudes of the employees of large corporations generally do not matter to the overall goals of the corporation. Members matter to the functioning of the group insofar as they perform many of its necessary processes, but in their roles as group members they are bound by the aims of the group agent. There is a separation between the attitudes of the members and the ‘attitude(s)’ of the group. In rare cases, the members can try to force some change in the group’s usual activity. For instance, in the case of a strike, much of the material of the group refuses to pursue its needs, so it must

---

<sup>1</sup> ‘Individual’ here is related to ‘individuation’ rather than ‘individualism’ as social theorists understand the term. For those interested in discussions of individualism and holism, see List & Spiekermann (2013). The arguments presented here are compatible with ‘supervenience individualism’ as described there as well as with holism. This paper focuses on the general criteria of individuation for agents and how that applies to groups, which is distinct from problems concerning the relationship between individuals and collectives. Thank you to an anonymous reviewer for pushing me to clarify this point.

<sup>2</sup> All agents, then, are individuals in this sense. To avoid confusion, so-called ‘individual’ agents (humans, dogs, etc.) will here be called ‘singular’ agents. The most obvious distinction between the two is that group agents are physically discontinuous systems.

adapt to these new circumstances. Under normal conditions, however, the members are each alone and almost entirely constrained. It is only as a rebellion to this constraint that striking even makes sense at all – if the attitudes, beliefs, or goals of the members and that of the group were aligned, then there could never be a situation in which a majority of the group’s members go on strike. Even those in positions of power within these structures do not have a great deal of personal autonomy when they are acting on behalf of the group. It is illustrative of this point that when Steve Jobs died, nothing fundamentally changed for Apple. This is because the corporation was not bound by the man; he, instead, only served a particular function within the group structure. These are the sorts of groups that I consider the object of analysis here. These I will call *proper* or *genuine* groups, as opposed to *mere collectives*, which involve those ‘groups’ of people with shared intentions. Intuitively, groups that count as proper groups are things like corporations, political parties, NGOs, and universities.

Being able to differentiate between proper groups and mere collectives will better allow us to understand and address unjust collective actions, coordination problems, and will improve our theorising about social and political problems more generally. It matters, for instance, whether a particular injustice was perpetrated by a mere collective of singular agents or by a genuine group agent. In the former case, responsibility lies solely with the people involved. Preventing the same injustice from occurring again should therefore involve improving (moral) education for singular people, addressing factors that affect particular people’s lives, and other individualised responses. In the latter case, however, the group agent determines the best available actions, and so our resolution more likely lies in restructuring the agent or influencing its environmental incentives. Furthermore, by simply having a better picture of our social and political landscapes in terms of proper groups and mere collectives, we will be able to think more accurately about influences on our own ways of thinking and living together.

To provide an account of proper group agents and answer the titular question, I will argue for what I call here the *autonomous individuation* account of individuality found in the enactive literature on agency (Di Paolo & Thompson, 2014; Di Paolo et al., 2017). In one sense, then, this paper is a partial defence and expansion of the enactive theory of agency to groups. Hence, many of the views expounded and defended here, as well as the methodology of the argument more generally, may be distinctive of that approach. That said, I believe the extra work involved in expounding the atypical use of certain concepts present in the enactive literature is worthwhile since the theory provides us with the most robust account of group agency presently available.

I begin the argument by explaining ‘agency’ as it is used here and defending the basic enactive definition of agency as involving a system acting in order to achieve some goal (Barandiaran et al., 2009, p. 369). This view, I argue, is the common understanding of agency in the social ontology literature, being either implied or stated as the initial perspective from which much reasoning on group agency begins. Therefore, it follows that individuality is a necessary element of defining group agency as it is commonly understood. Next, I will discuss the concept of individuality itself, explicating the necessary features of a definition of individuality by drawing on the

work of Barandiaran et al., (2009), Jonas (1966), and Meincke (2019). I then argue for the autonomous individuation account by showing how autonomy is central to individuality in agents and then providing the enactive definition of autonomy.

I will then compare the autonomous individuation account with the other candidates for definitions of individuality in the group agency literature, focusing on the work of Rovane (2019) and Tuomela (2013). I demonstrate that their definitions often fall short in many respects compared to the autonomous individuation account. I will conclude by considering a few examples of different kinds of groups in order to show that the autonomous individuation account of individuality is a robust and operationalizable account, and to demonstrate how it differs from the other available accounts.

## 2 Agency: an overview

Agency, as it is understood here, refers to ‘at least, *a system doing something by itself according to certain goals or norms within a specific environment*’ (Barandiaran et al., 2009, p. 369). This basic definition comes from an investigation of the discussions on agency in cognitive science and adaptive behaviour modelling and involves three essential parts (Barandiaran et al., 2009, pp. 368–9). First, there must be a system that is separate from its environment. This is the individuality criterion, with which we are primarily concerned here. The second and third criteria concern the individual’s ability to *act* and the *goals* or *norms* according to which that system acts. These are called the *interactional asymmetry* and *normativity* criteria respectively (Barandiaran et al., 2009, pp. 369–72). I take this to be the basic definition of the kind of agency that I am interested in here. ‘Agency’ is used in different ways across (and even within) disciplines, such as in chemistry, sociology, or meta-ethics.<sup>3</sup> When I claim that groups can be agents, however, I mean that they can be agents in the same way that humans, dogs, and other organisms can be agents.<sup>4</sup> What distinguishes different agents of this sort will be in the particular ways that their agency is established or manifested. A human agent, for instance, can consciously determine their own goals and can reflect on the best ways to influence the world to achieve those goals. A bacterium, on the other hand, is much more restricted in the norms that it could possibly pursue and is unlikely to be capable of any kind of conscious reflection on its reasons for taking one action over another. The most obvious difference between group and singular agents concerns the internal relations between their parts. Group agents are physically discontinuous systems whereas singular agents are physically continuous. Nevertheless, whether the system is continuous or discontinuous, there must be something that makes it a ‘system’.

This basic idea of agency is notably similar to other ideas expressed in the group agency literature, which is indicative of the fact that the core concept of ‘agency’

<sup>3</sup> Thank you to an anonymous reviewer for raising this point.

<sup>4</sup> Even single-celled organisms and plants appear to satisfy the three criteria for agency insofar as there is an identifiable structure that influences its environment in order to, at least, persist. This is a common view for enactive theorists of agency (Barandiaran et al., 2009, p. 374), and it is one that I follow them in holding.

being theorised is the same for both the enactivists and for social ontologists concerned with group agency. List and Philip Pettit's (2011) definition, for instance, differs predominantly due to a couple of additions. They argue that an agent is a system with representational states and motivational states capable of processing these states and acting on their environment in order to pursue their motivations (List & Pettit, 2011, p. 20). The addition of representational states and the capacity for processing one's states ultimately suggests that agency necessarily involves a particular cognitive framework, but the core idea that there is still some distinct system or individual doing something for a goal remains.

Tuomela (2013) is similarly explicit: 'The account [Tuomela's] regards organized groups that are capable of action as functional group agents' (p. 13); 'the notion of group agent (or that of a group capable of action)' (p. 46). This capacity for action depends on the singular agents who make up the group acting together for the same authoritative group reasons (Tuomela, 2013, p. 23). In this way, the group constitutes a system that is evidently distinct from its environment acting according to its own, internally determined group reasons.

Finally, Tollefsen (2002), in arguing that group agents are intentional agents, does so on the basis that 'our explanations of the actions of organizations in terms of their beliefs, intentions, and desires are successful' (p. 397). So, again, an agent must be a system (an organisation) whose beliefs, intentions, and desires we are trying to explain with reference to the acts they have performed. Though she does not appear to have a preferred definition of how a system is constituted, Tollefsen does note that, to be a group agent, groups must in some way form a coherent whole: 'The performance of joint actions on the basis of group ends, shared intentions, joint commitments, or we-intentions might very well be the way in which corporate agents form and sustain their agency over time' (Tollefsen, 2015, p. 47). There must, she suggests, be some persistent entity that is in some way unified.

The idea that agency involves *at least* a system doing something in its environment to achieve a goal is evidently uncontroversial. Some accounts of group agency add additional criteria, as with List and Pettit, and some focus more heavily on particular aspects, as Tollefsen focuses primarily on the goals or norms of groups. Nevertheless, individuality, interactional asymmetry, and normativity are common to all. It is, I contend, the first of these that is most often overlooked, despite being a necessary condition of agency. Let us, then, give it the attention it deserves.

### 3 The conditions of individuality

Allowing us to identify group agents is an important part of any successful definition of group individuality, so here I will explicate the conditions of one. Following Jonas, Barandiaran et al., (2009) point out that an agential system must be capable of distinguishing *itself* as an individual and, in doing so, defining its environment for itself (p. 370). This is a given so long as we take agency to be an objective fact of certain systems, which is the non-metaphorical position taken by many other philosophers concerned with group agency (List, 2021, p. 4; List & Pettit 2011, pp. 2–6; Pauer-Studer 2014; Rovane, 2019, p. 4870; Tollefsen 2002, p. 396; Tuomela 2013, p. 47). We can-

not, then, impose individuality on agential systems. By virtue of being the locus of activity, the agent is necessarily a self-distinguishing system. This is why autonomy is central to understanding agency, as I will argue further in the next section.

That agents define their own identities as individual systems follows from the fact that agents are agents regardless of outside observers judging them so. But not all systems are genuine systems without external observers (Barandiaran et al., 2009, p. 369). This means that, when considering what constitutes a proper group, we cannot just assume that anything we describe or perceive as a ‘system’ qualifies. For instance, what belongs to a workspace as a system depends entirely on the functionality of the various parts in that space for the observer (Barandiaran et al., 2009, p. 369). This is a case where there is something that might be considered a ‘system’ according to certain understandings of the term, but this is only because of its use-value to external observers rather than it being a self-individuating system. We will need a more specific understanding of what constitutes the kind of systems we are interested in.

Furthermore, List & Pettit (2011) claim that group agents must be able to persist through changes in membership and that ‘any multi-member agent must be identifiable over time by the way its beliefs and desires evolve’ (p. 32). The thrust of these claims is correct, though they require some amendment. The first of their points, that group agents must be able to persist through changes in membership, is reminiscent of Jonas’s (1966) argument that organisms cannot be identical to their material parts. Jonas (1966) argues that if we were to take a purely material picture of the world, ‘all the features of a self-related autonomous entity would, in the end, appear as purely phenomenal, that is, fictitious’ (p. 78). Agential systems, for Jonas as for List and Pettit, are dependent for their existence on the availability of material parts while at the same time maintaining a separate functional identity that is not the same as the identity of its material parts.<sup>5</sup> Jonas (1966) calls this relationship one of ‘*needful freedom*’ (p. 80): the agent both needs and is free from its matter. Just as a person’s identity does not change while they breathe, eat, and sweat, so too does a proper group maintain its identity through changes in membership.

It is for this reason that individuality refers to the individuation of agents only, which is a special case of individuation. As Wayne Christensen and Mark Bickhard (2002) have aptly pointed out, there are a number of properties that can serve as observer-independent criteria for identifying a given system (p. 8). Physical cohesion is one such property. It allows certain systems to be individuated from their environment insofar as they are causally bonded in particular ways. If you kick a small rock, the entire rock will move while the ground below it will remain in place (Christensen & Bickhard, 2002, p. 8). In this case, however, the entire identity of the rock is given by its physical cohesion. If these particular physical bonds are broken, the rock no longer exists. Agents, on the other hand, actively maintain their structure by taking in new material to replace what has been or will be lost. This is true even in singular

---

<sup>5</sup> The organism ‘introduced the tension of “to be or not to be” into the neutral assuredness of existence. It did so by assuming a position of hazardous independence from the very matter which is yet indispensable for its being; by divorcing its own identity from that of its temporary stuff, through which it is yet part of the common physical world’ (Jonas, 1966, p. 4).

agents. Though singular agents are physically continuous beings, their parts are not permanently cohesive the way that a rock's parts are. For both singular and group agents, then, individuality cannot be defined in strictly physical terms.

Hence, it makes sense that List and Pettit point to the idea that an agent must be flexible over time. The latter part of their claim is that group agents must be identifiable over time *by the way* their beliefs and desires evolve. I do not agree with their specific claim, but it does point to a more general point that holds true for any agent. Meincke (2019) argues that things that persist through time must be conceived of as 'stabilised processes' and 'what matters from a metaphysical point of view is that any process of whatever kind persists as long as stabilisation can be maintained' (pp. 24–5). List and Pettit gesture toward the general rule that Meincke is concerned with – the persistence of identity over time. While the rock persists for as long as its molecular bonds hold, agents persist for as long as they can continue to be active systems. So, group agents may not be identifiable *by the way* they evolve, since we do not know without a picture of the system just what is evolving; but it is necessarily the case that persistence, as Meincke points out, requires change. Hence, it is a general rule that group agents do need to 'evolve' over time in order to persist. This might occur very quickly or exceptionally slowly, but in the face of a changing environment, the agent needs to adapt or the external conditions for its survival will no longer be met. Blockbuster as compared to Netflix serves as an apt and familiar example here. Beliefs and desires are not core to this picture, instead what is necessary is an evolution in behaviour. This may or may not result from evolving beliefs and desires, but we need not take any position on this particular claim. Taking these points together, the other goal of a definition of individuality is to define the agent's conditions of stabilisation over and above a particular relationship between its material components.

At the same time, however, it is important that we are able to distinguish between the agent itself and those parts of its environment that it relies on for its stabilisation. Again, organisms require food and water, but the sources of these things in the environment of the agent are not themselves parts of the agent. Businesses similarly require customers, but the customers are not constitutive of the business itself. In both cases, we must be able to distinguish between those parts of the world that belong to and constitute the agent and those parts that are external to it that are nevertheless necessary for its persistence.

#### 4 The autonomous individuation account

The enactive definition of autonomy provides, as Di Paolo and Evan Thompson (2014) put it, the criteria for the self-individuation of bodies (p. 69), where a body is *not* 'constituted exclusively by its biochemical or physiological processes' (p. 72). Autonomy here still refers at the broadest level to self-governance (see also Barandiaran & Egbert 2013, p. 8; Barandiaran 2017, p. 410; Christensen & Bickhard 2002, p. 3). Because agents necessarily demarcate their own boundaries and define their own environments, as I argued above, autonomy is vital to understanding agency. The claim being made here is that it is precisely this self-governance that generates the

agent's individuality.<sup>6</sup> Being an autonomous system just is what sets the agent apart from its material components and the rest of the physical world. How exactly this occurs is the point of the somewhat technical definition given below.

There are two aspects to autonomy: *operational closure* and *precariousness* (Di Paolo & Thompson, 2014, p. 69). The first, operational closure, is of primary interest, though precariousness is an important addition (Di Paolo & Thompson, 2014, p. 72; Di Paolo et al., 2017, p. 116).

For a system to be operationally closed, its constitutive processes must collectively actively produce and sustain those self-same processes (Di Paolo et al., 2017, p. 112). So, if process A sustains process B, which sustains process C, which sustains process A, then the system ABC is operationally closed. Think of the ways a plant's roots, stems, and leaves all sustain themselves and each other. Here, we can see that the parts of the system are related to the system as a whole via their production and sustenance of that self-same system, which is stabilised by this active producing and sustaining.

Of course, being self-sustaining does not imply that an operationally closed system is cut off from its environment. There are a few ways in which other processes might influence the system without thereby being a part of the system. Furthermore, it is through the systems' interactions with these external processes that it defines its environment, which is crucial for both a metaphysical understanding of group agents and for pragmatic reasons concerning how we might influence groups by changing their environments.

First of all, some processes enable the system in question while not themselves being enabled by the system (Di Paolo & Thompson, 2014, p. 71; Di Paolo et al., 2017, p. 114). In plants, the sun acts as an enabling condition for their photosynthesis, while not being sustained by plant life itself (Di Paolo & Thompson, 2014, p. 71). Likewise, for groups we could consider the conditions of capitalism broadly (private property, private control of the means of production, the legal structures that protect and ratify these, and so on) as enabling conditions for certain corporations. Furthermore, there are processes that act as boundaries and constraints for operationally closed systems (Di Paolo et al., 2017, p. 114). These can be the same processes that enable the system in the first place. There may, however, also be boundaries that are not created by the enabling conditions for that system, as minimum wage laws and the threat of union action are against modern corporations. Note, as Di Paolo et al. point out, that these enabling and binding conditions are implied by the organisation of the system in question (Di Paolo et al., 2017, p. 114). The sun is a necessary enabling condition for the plant, as determined by the organisation of processes in the plant. We, as external observers, may or may not notice this enabling condition, but it exists regardless. In this sense, operational closure generates individuality for the agent and distinguishes it from its environment while giving us as external observers a useful tool for uncovering the structures of the agential system in a non-arbitrary manner.

---

<sup>6</sup> There can, of course, be other definitions of autonomy that are employed in different contexts. The autonomy that is being defined here should be understood strictly in the context of the understanding agency and agents.



Next, a word on precariousness. By precariousness is meant: ‘in the absence of the enabling relations established by the operationally closed network, a process belonging to the network will stop or run down’ (Di Paolo & Thompson, 2014, p. 72). So, all of the processes that are a part of a system are precarious because they are actively enabled by the other processes in that system. This concept is necessary to avoid the inclusion of trivial cases in the set of self-individuating systems. A crystal, for example, satisfies the conditions of operational closure since ‘chemical interactions lead to the spontaneous growth of a clearly identifiable entity, which thereafter is maintained over time’ (Di Paolo et al., 2017, p. 116). The issue is that this crystal does not have to *do* anything after being formed to maintain its existence. It simply persists, without its processes needing to *actively* maintain each other.

That said, the idea of precariousness is not as vital to the account presented here as operational closure. There is in theory no issue for a definition of agency that a crystal counts as self-individuating, since the idea that an agent must be an active entity will be taken care of by the concepts of interactional asymmetry and normativity. All agents need to be self-individuating systems, but not all self-individuating systems need to be agents, since there are other criteria that must be satisfied. Whether the crystal is a self-individuated system or not is neutral to the project of defining group agents. Nevertheless, it is reasonable to think that all agents really are precarious in the sense argued for above. Organisms must actively sustain themselves. Corporations and political parties must similarly be engaged in constant activity or else they will fail.

The autonomous individuation account suggests that agents are individuals because they autonomously make themselves such via their active determination and maintenance of their own structures. Their precariousness, furthermore, produces the normative demands with which we, as biological systems, are intimately familiar. What they are and what they do, then, are self-governed.

## 5 Rovane and Tuomela

Here, I will compare the autonomous individuation account with two of the more substantial engagements with the individuality criterion, namely Rovane’s (2019) and Tuomela’s (2013). I will argue that each of their accounts face issues that the autonomous individuation account manages to avoid.

Rovane (2019) sums up her account of agency thus: ‘*wherever there is a commitment to meeting the normative requirements that define individual rationality there is an individual agent*’ (p. 4874). The normative requirements in question are consistency, closure, and transitivity (Rovane, 2019, p. 4873). To be consistent, an agent must resolve conflicting beliefs. Closure is achieved by accepting the implications of one’s beliefs. Transitivity is the ordering of preferences. A group agent exists just in case there is a collection of people who together meet the normative requirements that define individual rationality (Rovane, 2019, p. 4870). According to Rovane’s definition of agency, the individuality criterion is satisfied by the presence of the relevant commitment. This commitment defines the boundaries and sustained existence of the group since only those parts of the world that are teleologically oriented toward

this commitment to the norms of rationality can possibly count as a part of the group, and the group persists so long as the commitment remains. So, to determine whether a particular group is a proper group, we ask whether that group has consistent beliefs that it at least attempts to follow through on and whether it prioritises potential tasks in terms of greater or lesser importance. Here, we will likely look for organisational structures that allow for the relevant processes to occur. This might take the form of a voting procedure, or there might be individuals authorised to make decisions on behalf of the group due to their (purported) expertise.

Importantly, Rovane's (2019) definition of agency is grounded in *human* agency, both in the sense that groups always seem to be made up of humans (p. 4870) and that groups are agents because they are relevantly like humans (p. 4877). The first assumption is not a significant issue, since in that paper Rovane is concerned with the question of whether or not group agency is a social phenomenon, and so she may argue that she is concerned with groups made up of humans in this particular context. Still, it is worth noting that non-human animals like bees and ants most likely form group agents just like humans do and excluding those groups as counting without justification limits the account.

The second assumption, however, is likely to lead to an inflated conception of agency where agency is taken to involve some number of properties that do not belong, especially when defining non-human agents such as group agents. In Rovane's (2019) case specifically, these are mental properties, including intent (p. 4871), beliefs, attitudes, and commitment (p. 4873).

Agents are not necessarily mental beings. The enactive theory of agency defines agents without needing to rely on the concepts Rovane employs, which implies that using mental concepts to define agency in general requires some justification. The normativity criterion might be the primary avenue through which mental features make their way into defining agency. Some might take the possession of norms and goals, for instance, to imply that the agent must have the capacity to reflect on these norms or to choose their goals. This, however, is mistaken. The goal of survival among biological systems is evident and is one that does not require a mind to adopt. Plants and amoebae, neither of which is often taken to have minds capable of beliefs or attitudes, actively pursue their own survival. The lack of indifference to the conditions a system finds itself in is a definitive characteristic of the living system (Cagnu-ilhem, 1991, p. 126). Hence, as anyone who has had a plant by a window will know, plants will orient themselves and their leaves toward the sun. This is just to say that mindedness and agency are, at the very least, conceptually separable and, therefore, it cannot simply be assumed that an agent has a mind or mental capacities. If it must, then this is something that needs to be argued for, not merely assumed.

For what it's worth, I find the idea, mentioned by David Spurrett (2020) on Twitter (of all places), that '[c]ognition is the control of agency' far more likely. This would help to explain the apparent differences in complexity among agents. I, as a human, can control my actions, develop my own goals, rationally reflect on the best ways to achieve those goals, and so on. A bacterium, on the other hand, will not form goals that fall outside of its biological needs, and cannot reflect on the best ways of achieving its goals, though is still an agent since it is an individual that acts on its environment to achieve its own normative ends.

Rovane's argument in particular appears to be an instance of what Fred Adams and Kenneth Aizawa (2009) term the *coupling-constitution fallacy* (p. 81). This occurs when someone moves from the observation that process X is causally connected to process Y to the view that X is constitutive of Y (Adams & Aizawa, 2009, p. 81). In Rovane's case, her focus on human agents, in whom our cognitive processes are very likely causally connected with our agency, has possibly led to the implicit view that certain cognitive processes are constitutive of agency itself.

Because Rovane's definition depends on expanding an understanding of human agency to agents that are not themselves humans, we are left with two options: (1) her account implies certain attributes that require mental capacities that groups almost certainly do not have, or (2) her use of those terms is purely metaphorical. The first option implies we should reject Rovane's view. Even if singular humans can be said to have beliefs and preferences on behalf of the group, these are not the group's beliefs or preferences. Similarly, it is not clear how 'commitment' to the standards of rationality figures as a sturdy definition of the group's individuality. Rovane (2004) talks of a commitment to a group in terms of a person's conscious choices or feelings regarding the activities of that group (p. 194). This does not address the differences between the necessary processes that make up the group itself and external processes that the group relies on. Thus, it does not allow for a demarcation between genuine members of the group, such as people performing necessary tasks, and interested parties, such as politically engaged individuals who try to convince their friends to vote for their preferred party. The autonomous individuation account, on the other hand, does address these concerns since the demarcation of the agent's own parts from its external environment is built into the enactive definition of autonomy. It is, therefore, preferable for its greater clarity.

The second option is less serious but does suggest that if we want a complete understanding of the ontology of group agents – and for the sake of more technical work – we should figure out what 'beliefs', 'preferences', and 'commitment' are metaphors for. Here we might simply substitute a more robust account of agency, such as the enactive theory. Talk of mental states might be useful if we are just trying to give non-experts a rough overview of the general ideas, since the terms will more easily communicate the approximate idea without having to explain technical terms like 'operational closure' and 'precariousness'. If we are after an accurate, operational account, however, then technical terminology should not be a barrier.

For Tuomela (2013), a group agent is a mind-dependent entity with both fictitious properties and real causal powers (p. 47). They are partly fictitious in the sense that group agents do not really have intentional features that, Tuomela (2013) claims, depend on a biological brain such as the capacity for reason (p. 48). On the other hand, they have objectively real causal powers insofar as the existence of a group agent for its members produces certain outcomes in virtue of those members acting as group members (Tuomela, 2013, p. 47). Given the ontological mind-dependence of group agents for Tuomela (2013), he also argues that they must be collectively constructed and collectively accepted (p. 47). He offers the example of John and Jane, who jointly intend to paint their house together, to clarify his point (Tuomela, 2013, p. 49):

Now consider the dyad, a group agent, consisting of John and Jane. This group agent is collectively constructed. Simplistically put, John and Jane form a group agent because they (and others) take them to form a group. This view... is ontologically grounded by John's and Jane's relational state of joint intention (i.e., the individual we-intentions and the mutual awareness that it is ontologically composed of) and by their joint action dispositions.

The construction that is mentioned here involves a collective acceptance that is dependent on people's imaginative capacities (Tuomela, 2013, p. 49). The dyad is a group agent precisely because John and Jane (and others) imagine and accept that they form a group agent. They hence each adopt a we-intention which, roughly put, involves intending to play one's part in 'our' action (p. 78). This, then, allows the people to act intentionally together, and hence as a group. Although Tuomela believes groups themselves cannot have intentions, the *members* of the group can still intend to act collectively.<sup>7</sup>

Membership in a group agent involves (collectively) accepting the group's *ethos*. The group's ethos is 'the group's central, typically action-related constitutive properties' (Tuomela, 2013, p. 26). Painting a house, for instance, is a central, action-related constitutive property of the John and Jane dyad. To be a member of the dyad, John and Jane must each accept this ethos and the various actions related to it. New members, to act as group members, must also accept the group ethos. Furthermore, if the group is faced with a choice between courses of action in relation to achieving a group goal, this choice is determined by the members of the group collectively accepting an attitude, which becomes the group's attitude (Tuomela, 2013, p. 123). Hence, on this account, a group agent accepts  $p$  as true if and only if its members collectively accept  $p$  as true for the group (Tuomela, 2013, p. 127). The members of a group collectively accept  $p$  as true if and only if they jointly have an attitude expressed by  $p$  that is of use to the group, by which he means it promotes the group agent's goals (pp. 127-8).

Finally, it is worth mentioning that, while for Tuomela (2013) the group agent ontologically depends on its members, they can be considered 'position-holders', meaning one person may leave and another can come along to take their position without this changing the identity or character of the group agent (p. 26). For instance, the John and Jane dyad is a group agent that has an ethos related to painting their house. Now let's say Jean joins John and Jane, endorsing the group ethos and we-intending to paint the house. By taking up these psychological attitudes, Jean becomes a member of the group. If John leaves the group but Jane and Jean continue to promote the group's ethos, then the group itself will persist. Hence, the group's identity is not directly dependent on its members, although its continued existence does depend on having members who fulfill the relevant roles. Along similar lines, he maintains that exactly the same people can form multiple group agents. The groups will differ insofar as they differ in ethos and activity (Tuomela, 2013, p. 49). John and Jane can have a house painting group, a book club, and a band, all of which will

<sup>7</sup> Thank you to an anonymous reviewer for highlighting this important aspect of Tuomela's account of agency.

be separate group agents despite containing the same singular human agents. The groups have their own identities, which are initially determined by their founders, but thereafter become their own.

Given the notions of ethos and collective acceptance, Tuomela has greatly overestimated the extent of the knowledge, power, and agreement that is necessary among group members, even those we would consider executive members in a non-egalitarian group. As Jonas (1966) argues, there is a difference between *having* and *servicing* a purpose (p. 122). I may form a goal – a purpose – and then carve it up into disparate pieces to be performed by a number of other singular agents who, while knowing their own goals, know nothing of the broader context in which they operate; they are ‘goal-blind’ (Jonas, 1966, p. 123). Likewise, a group might demand certain actions by various means of its members who are each goal-blind but who together serve the overall purpose of the group. This demand might come in the form of felt pressures. This pressure can come from members enforcing the rules, or even from non-members who expect you to play your part given what they know about the group itself. It might come in the form of psychological pressure to perform the duty you have committed yourself to and, perhaps, from fear of failure. It might come from economic or political circumstances. Sources of coercion abound, even in our own minds. In short, there need not be anything like a collective acceptance of  $p$  in order for there to be a group acceptance of  $p$ .

To push the point yet further, there need not even be *members*. Again, as Jonas (1966, p. 123) explains:

I can even reduce the steps to such primitive elements that I can dispense with human agents altogether. It is precisely this dissociability of purpose and execution which permits us to delegate the latter so extensively and distributively to others, to whole chains of subagents, and even to machines.

The kinds of structures with which we are concerned when we think about group agents are, in some cases, already made up substantially of automata and automatic processes – machines to make parts and machines to fix the machines that make the parts; software running websites, placing ads, and collecting data; self-serve checkouts that turn the customer into their own cashiers and cashiers into dual-role IT and security personnel. And yet, for Tuomela, these parts are invisible. This is understandable if we are concerned with human agents and their interactions and roles in groups. It is not, however, a viable position to hold in light of current and emerging group agents and their increasingly automated functions. To offer an account of group agency that misses these functions is to describe the human body without skin or fat or bones. In doing so, we run the risk of developing an almost instantly outdated concept of group agency.

The autonomous individuation account, in contrast, better captures the automated processes of an agent within the bounds of that agent precisely because it is agnostic about the particular material constitution of the agent. Tuomela does not ask what it is to be individuated, but rather moves directly to an account of how people can come together to form a supposedly individual group agent. The enactivists focus instead on the more fundamental question of individuality, remaining neutral on material

constitution. It is precisely this methodological difference and the agnosticism of the autonomous individuation account toward material constitution that allows for the accurate employment of it in the contemporary world. Many groups are at least partly automated systems and are extremely large, such that it is extremely unlikely that all (or even many, in some cases) of the group's members can come close to collectively accepting the group's ethos.

The autonomous individuation account allows us to accurately demarcate the boundaries of autonomous groups while Tuomela's account would be better framed as concerning the interactions between groups and their human members. If we adopt this perspective, then we need not reject his view, but only to reconfigure our understanding of it. In any case, as a definition of individuality, the autonomous individuation account fares better.

## 6 Identifying group agents

Examples of supposed group agents abound. Here, I will consider some of the kinds of groups that other philosophers of group agency have suggested in order to show both the differences in views and as a proof of concept for the implementation of the autonomous account of individuality. The differences discussed here are relevant since they matter for thinking about moral and legal responsibility, power relations and structures, and any other practical matters pertaining to autonomous group agents. If a friend group commits a crime, we can work out how culpable each individual was and which actions they took and deal with them each accordingly. If a corporation commits a crime, we might also punish some number of key individuals, but the group itself must also be dealt with in some way. Responsible individuals in the group agent case might also get off a little more lightly because of coercive forces within the structure of the group. It might even be that no individuals at all suffer the full brunt of legal force. In a mere collective, if any individuals were coerced, they must have been coerced by another individual, and hence there will always be at least one human being that we can hold fully responsible. This is, of course, just an outline of how the differences in what we consider a group agent matter. How exactly group agents should be held responsible, how their members should be held responsible, and so on, all warrant their own discussions. It is worth pointing out here, however, so that the weight of the following discussion is more obvious.

Recall first Tuomela's (2013) dyad consisting of John and Jane, who jointly intend to paint their house together. Tuomela (2013) claims that, simply put, 'John and Jane form a group because they (and others) take them to form a group' (p. 49). I have argued above against Tuomela's account of group agency. Here, I intend to discuss the example to highlight the differences between the autonomous individuation account and Tuomela's account.

In Tuomela's example, the dyad's collective activity is to paint a house. This activity is established by John's and Jane's *we-intentions* to paint the house together. For the dyad to become a group agent on the autonomous individuation account, their forming this agreement together would have to render the dyad itself an autonomous system, now in a relationship of needful freedom from John and Jane. But the group

is neither operationally closed nor is it precarious. If John, due to an injury say, can no longer uphold his end of the bargain, Jane is still capable of continuing on with the painting of the house. Her activity is not sustained by John's activity; hence the dyad is not operationally closed. Since her activity does not depend on John's, the dyad cannot be precarious.

This does not mean there are no dyadic group agents. It just means there are further conditions that need to be met than the couple simply agreeing that they form a group. If John and Jane formed a company that paints houses, for instance, we might have a group agent on our hands. In the case that they form a company, the company itself generates an impetus toward profit, such that John and Jane as people are not even strictly necessary. Of course, the formation of the company and its getting off the ground requires their intentional activity, but once it is off the ground it certainly might qualify for group agency should the conditions of operational closure and precariousness be met. We could easily say that these conditions have been met should it be possible for their company to persist without John and Jane – if they have hired other individuals who will continue the activity of the house painting business even if John and Jane are not involved. It might also satisfy these requirements if John and Jane each take up different essential roles that rely on each other. For instance, if John takes up the tasks of finding jobs, giving quotes, sorting out where and when houses need to be painted, and so on, and Jane does the actual painting, then we have, as Tuomela would put it, John and Jane as position-holders in a group agent that, for now, happens to be a dyad. If John stops booking jobs, Jane will have nothing to paint, the group will make no money, and they will go out of business. If Jane stops painting, then even if John books jobs, they will not be fulfilled, and the same will happen. This, however, is only the case insofar as neither bothers with hiring a replacement. Since they are both only position-holders, fulfilling particular functions that the group agent demands given the kind of group agent it is (namely, a house painting business), they are both also replaceable. Once again, the issue with Tuomela's account is his emphasis on the singular agents who make up the group, rather than on the group agent itself.

In light of this discussion, it is clear that there is a rather thin line between singular agency in mere collectives and group agency. What is less obvious, but is equally worth remarking on, is the minor difference between group agency and singular agency without a collective. In the case where John quits and Jane does not hire someone new, but instead takes on John's roles for herself on top of her painting duties, the group agent disintegrates and we are left with a singular agent, Jane, who is monetising certain abilities she has. What is special about group agents is their physical discontinuity, where their functions are performed by separate entities who are not tied together spatially or temporally. When the functions that might usually be performed by a group agent are instead performed by a single person, there is no longer any point in thinking about the business in terms of group agency, since it loses all the distinctive features of groups, both ontologically and from a pragmatic perspective; the results of thinking about responsibility and group agency are irrelevant, since there is only one person acting and making decisions.

Next is an example that List & Pettit (2011) take not to constitute a group agent. In their scenario, there is a swimmer struggling in the water at the beach and a number

of people notice the swimmer's plight. These people together form a chain so that a lifebelt can be thrown to the swimmer without putting anyone else in danger (List & Pettit, 2011, p. 34). For List & Pettit (2011), the group of rescuers fails to form a group agent because it does not have 'a single system of belief and desire' (p. 34). If we take 'belief and desire' to refer to particular characteristic kinds of activity, then List and Pettit are mistaken that the chain of rescuers does not have a system of belief and desire. The chain of rescuers is normatively oriented toward saving the drowning swimmer, as is evident from the chain's activity and its reason for being formed. If they instead mean that the individuals who make up the chain do not have the relevant beliefs and desires together, then the claim is not relevant when thinking about group agency. I have already argued this point in relation to Tuomela – a group agent may easily be made up entirely or almost entirely of goal-blind singular agents. So long as they perform the relevant functions on behalf of the group, the group itself will persist. Finally, if they mean that the group itself should have the mental properties of belief and desire, then this is again mistaken. As argued above in the discussion on Rovane, it is a mistake to assume that agency and mindedness go together, and hence the idea that an entity should be excluded from the possibility of agency because it lacks such minded qualities as belief and desire is incorrect.

Furthermore, we should recall that for List & Pettit (2011) groups, to count as agents, must persist in some sense. They argue that the group cannot be an agent because, due to its failure to form a single system of belief and desire, we cannot predict what it will do in the future. Group agents must indeed persist – they must be stabilised for a certain period of time. But just how long they must persist for is not relevant. There is, then, a rather arbitrary understanding of persistence that is implied by List and Pettit's argument. Different groups – different agents in general – will have different characteristic timescales.

The group of individuals forming the chain clearly form an operationally closed and precarious system. The chain relies on all of its members for it to be formed and sustained. If the strength of someone in the middle of the chain fails, then the chain breaks and half of the group members will be in danger. Likewise, if someone refuses to participate, the chain is shorter and the job harder – though in this case the group does not necessarily collapse but is instead different in character since the character of this group is given largely by the length of the chain and, hence, its ease in saving the struggling swimmer. The fact that after the swimmer has been saved the various participants will go their different ways, thereby destroying the group, does not matter. The group exists as an individual for that particular task.

Taking both of the examples just discussed together, the autonomous individuation account likely implies that there are less group agents than Tuomela's account does, but more than List and Pettit's. List and Pettit's account denies that John and Jane are a group agent, as the autonomous individuation account does. On the other hand, Tuomela's account coincides with the autonomous individuation account in accepting that the chain of rescuers is a group agent. However, though there are conclusions in common between the autonomous individuation account and the accounts of other philosophers, they come to their conclusions for the wrong reasons. I have argued here for a robust, operational account of group individuality that is able to distinguish between mere collectives and proper groups in a non-arbitrary manner. Agency is not



dependent on an arbitrary amount of temporal persistence, nor is it simply based on accepting that something is an agent. Hence, although there is some overlap between the account endorsed here and its philosophical competitors, it is ultimately the differences in methodology that matter.

Finally, I will consider one case on which List & Pettit (2011) and I agree: the contemporary university (p. 194). Universities, being much larger than Tuomela's dyad or List and Pettit's human chain, are harder to provide a full account of in terms of their autonomous individuation, though a basic overview should be sufficient to get the point across.

Teaching and research are the most obvious processes of most universities.<sup>8</sup> A university, then, relies on its academic staff to fulfil these processes in order that its functions be performed. The academic staff, however, cannot properly do their jobs – certainly not as members of the university they belong to – without administrative staff, management, and students to teach. Likewise, each of these processes relies on each of the others. Students need to be able to enrol in courses, staff need to get paid, plans in case of emergency need to be formulated and acted upon where necessary, among a myriad of other processes dutifully performed by the people who make up the university in question. A university, then, constitutes an operationally closed and precarious system. All of its processes rely on and maintain other processes in that same system, and should one of the processes stop, the whole thing will quickly begin to run down. Therefore, universities are autonomously individuated systems. This conclusion should be uncontroversial, and that is precisely the point. The autonomous individuation account differs from alternative accounts when considering edge cases but, as I have just shown, quite easily handles obvious cases.

What makes all of the above examples group agents (or not, in the case of John and Jane), is simply that there is some individual system into which people (and machines) can be fitted to perform the vital functions of the system. The notion of group agency and group individuality defended here does not depend on any particular legal or other social conditions being met. Human group agents and beehives likely both have different social conditions that need to be met for a proper group to form. These social conditions will be determined by the kinds of beings that make up the group. This will be relevant in analyses of particular groups but does not change the concept of group agency itself. Legal requirements similarly do not impact on the concept of group agency, though they do form a part of the environment for legally operated groups. That there are illegal groups, such as drug cartels, is indicative of the fact that these legal requirements do not impact on the ontological possibility of genuine group formation.

---

<sup>8</sup> There are colleges, such as All Souls College, Oxford and research institutions like the Desert Research Institute in Las Vegas that function a bit differently to a typical university – All Souls College has no undergraduate students, for example. These kinds of institutions are still likely to satisfy the requirements of operational closure and precariousness but would require their own analysis. Thank you to an anonymous reviewer for bringing these examples to my attention.

## 7 Final remarks

Defining and understanding the individuality of group agents is, I have argued, an often-overlooked issue. This is despite their being of fundamental importance to an accurate assessment of the world and to political practice. Here, I have shown that for a group to be a group agent, it must be an autonomous group in the sense that it is operationally closed and precarious. Its operational closure is dependent on its co-constitutive processes and their performance. That an agent is precarious expresses the fact that agents are active, adaptive systems.

Of course, even a complete definition of the individuality of group agents does not amount to a complete theory of their spatiotemporal relations, or their interactions with one another and with their members, and so on. These further discussions are necessary for understanding fully how group agents scaffold or constrain our interactions with each other, our politics, even our lives. To do all of this, however, it is crucial that we have an accurate picture of the systems that we are trying to understand and, perhaps, address.

**Acknowledgements** Thank you to Patrick McGivern and Nick Brancazio for their invaluable discussion, guidance, and feedback at every stage in the production of this paper. I would also like to thank my two anonymous reviewers for their very thoughtful comments and helpful feedback.

**Funding** Not applicable.

**Open Access funding enabled and organized by CAUL and its Member Institutions**

**Data Availability** Not applicable.

**Code Availability** Not applicable.

## Declarations

**Conflicts of interest/Competing interests** Not applicable.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adams, F., & Aizawa, K. (2008). Why the Mind Is Still in the Head. In P. Roberts & M. Aydede (Eds.), *The Cambridge Handbook of Situated Cognition* (pp. 78–95). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816826>

- Barandiaran, X. E., Di Paolo, E., & Rohde, M. (2009). Defining Agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Journal of Adaptive Behavior*, 17(5), 367–386
- Barandiaran, X. E., & Egbert, M. (2013). Norm-Establishing and Norm-Following in Autonomous Agency. *Artificial Life*, 20(1), 5–28
- Barandiaran, X. E. (2017). Autonomy and Enactivism: Towards a Theory of Sensorimotor Autonomous Agency. *Topoi*, 36(1), 409–430. <https://doi.org/10.1007/s11245-016-9365-4>
- Bratman, M. E. (1992). Shared Cooperative Agency. *The Philosophical Review*, 101(2), 327–341. <https://doi.org/10.2307/2185537>
- Canguilhem, G. (1991). In Fawcett (trans.) (Ed.), *The Normal and the Pathological*. C. R.
- Christensen, W., & Bickhard, M. (2002). The Process Dynamics of Normative Function. *The Monist*, 85(1), 3–28
- Di Paolo, E., Buhrmann, T., & Barandiaran, X. E. (2017). *Sensorimotor Life: An Enactive Proposal*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198786849.001.0001>
- Di Paolo, E., & Thompson, E. (2014). The Enactive Approach. In L. Shapiro (Ed.), *Routledge Handbooks in Philosophy: The Routledge Handbook of Social Cognition* (pp. 68–78). Routledge. <https://doi.org/10.4324/9781315775845>
- Gilbert, M. (2009). Shared intention and personal intentions. *Philosophical Studies*, 144(1), 167–187. <https://doi.org/10.1007/s11098-009-9372-z>
- Jonas, H. (1966). *The Phenomenon of Life: Toward a Philosophical Biology*. Northwestern University Press
- List, C., & Pettit, P. (2011). *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199591565.001.0001>
- List, C., & Spiekermann, K. (2013). Methodological Individualism and Holism in Political Science: A Reconciliation. *American Political Science Review*, 107(4), 629–643. <https://doi.org/10.1017/S0003055413000373>
- List, C. (2021). Group Agency and Artificial Intelligence. *Philosophy and Technology*, 1–30. <https://doi.org/10.1007/s13347-021-00454-7>
- Meincke, A. S. (2019). The Disappearance of Change: Towards a Process Account of Persistence. *International Journal of Philosophical Studies*, 27(1), 12–30. <https://doi.org/10.1080/09672559.2018.1548634>
- Pauer-Studer, H. (2014). A Constitutive Account of Group Agency. *Erkenntnis*, 79(9), 1623–1639. <https://doi.org/10.1007/s10670-014-9632-y>
- Rovane, C. (2004). What is an Agent? *Synthese*, 140(1/2), 181–198
- Rovane, C. (2019). Is group agency a social phenomenon? *Synthese*, 196(12), 4869–4898. <https://doi.org/10.1007/s11229-017-1384-1>
- Spurrett, D., & [@DoctorSpurt] (2020, October 13). *Cognition is the control of agency. (Agency can also be studied behaviourally.)* [Tweet]. Twitter. <https://twitter.com/DoctorSpurt/status/1315868726642847744>
- Stapleton, M., & Froese, T. (2015). Is Collective Agency a Coherent Idea? Considerations From the Enactive Theory of Agency. In C. Misselhorn (ed.), *Collective Agency and Cooperation in Natural and Artificial Systems: Explanation, Implementation and Simulation* (pp. 219–236). Springer. <https://doi.org/10.1007/978-3-319-15515-9>
- Tollefsen, D. (2002). Organizations as True Believers. *Journal of Social Philosophy*, 33(3), 395–410. <https://doi.org/10.1111/0047-2786.00149>
- Tollefsen, D. (2015). *Groups as Agents*. Polity
- Tuomela, R. (2013). *Social Ontology: Collective Intentionality and Group Agents*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199978267.001.0001>