[PENULTIMATE DRAFT- Please cite the published version]

Group Agents and the Phenomenology of Joint Action

Abstract

Recently, philosophers and scientists have expanded our understanding of the structure and neural mechanisms of joint action. But the phenomenology of joint action has only recently become a live topic for research.

One method of clarifying what is unique about the phenomenology of joint action is by considering the alternative perspective of agents subsumed in group action. By group action we mean instances of individual agents acting while embedded within a group agent, instead of with individual coordination. Paradigm examples are educational bureaucracies, corporations, and nation states. There is a phenomenological difference between agents whose actions are subsumed within a group action ("the university did X") as compared to agents who act jointly ("we did x"). Attending to this difference clarifies what is phenomenologically distinctive about joint action.

Appealing to an Aristotelian account of agency and to the metaphysical concept of weak emergence, we argue that what makes paradigmatic group action distinctive is the relative inaccessibility, un-revisability, and evaluative simplicity of the group agent's goal(s) from the perspective of individual agents. This suggests that a distinctive feature of joint agency is the maintenance of a greater sense of individual agency. Put simply, joint agency is often experienced as an enhancement of the individuals' agency precisely because our paradigmatic agential powers are extended intersubjectively as we act together. In contrast, group agency often involves a loss of the sense of agency, precisely because it is the emergent group agent that maintains the agential powers.

1. Introduction

Humans are social beings. Hence, human agency must be understood in terms of joint action. Philosophers (e.g. Bratman 1992, 1993, 2014; Heal 2005, 2013; Tuomela 2006) and scientists (e.g. Carpenter 2009; Chaminade et al. 2012; Meyer et al. 2016; Sebanz et al. 2006; Trevarthen 1980, 2011) have done much to expand our understanding of the structure and neural mechanisms of joint action. But the phenomenology of joint action has only recently become a live topic for research (Seemann 2009; Pacherie 2008, 2012, 2014; Dewey et al. 2014; Zaadnoordijk et al. 2015).

One method of exploring the phenomenology of joint action is to consider how our individual sense of agency differs from that of acting together. This method is both necessary and useful; however, it should be supplemented by considering joint action from, as it were, the other direction: by comparing it with the experience of agents subsumed in group action.

By group action we mean the collective action of multiple agents embedded within a group agent, like an institution (see List and Pettit 2011, Rovane 2014, Pauer-Studer 2014). We claim that attending to the phenomenological difference between group activity ("the university did X") and activity performed jointly by individuals ("we did x") clarifies what is phenomenologically distinctive about joint action.

Appealing to an Aristotelian account of agency and to the metaphysical concept of weak emergence, we argue that paradigmatic group action is distinctive because of the relative inaccessibility, un-revisability, and evaluative simplicity of the group agent's goal(s) from the perspective of individual agents. The ends that structure joint activity typically lack these features. Moreover, joint agency is paradigmatically experienced as an enhancement of the individuals' agency because our central agential powers are extended intersubjectively as we act together. In contrast, group agency often involves a loss of the sense of agency because the emergent group agent maintains the agential powers.

In section 2, we examine what group agency and joint agency are. In section 3, we develop a phenomenological account of "what it is like" to constitute a group agent. In section 4, we explain these experiences by appealing to the conflict between the Humean conception of agency that applies to group agents and the Aristotelian notion of agency that applies to individuals, as well as to the concept of weak emergence. We tie these threads together in section 5, highlighting how these considerations illuminate the phenomenology of joint agency. We conclude, in section 6, by suggesting that the healthiest group agents are those that mimic the conditions of joint agency among their constituting members.

Finally, we should note that in developing our account, we rely on some disputed claims. We hope to show, however, that adopting the views presented here allows for a persuasive explanation of otherwise puzzling phenomena, and revealing the explanatory power, interest, and fruitfulness of a view is itself a form of argument.

2. Group Agents, Joint Action, and the Phenomenology of Collective Activity

For our purposes, we need not spell out necessary or sufficient conditions for the two categories, *group agency* and *joint activity*. The two kinds probably exist on a spectrum with fuzzy boundaries and edge cases.[1] Nonetheless, we take the distinction to be pre-theoretical and worthy of explanation. To begin theorizing about these phenomena, we define the categories through *paradigm cases*. A paradigm case is a recognizably central instance in reference to which other cases will either be included or excluded from a category by comparison. For example, we take the bureaucratic institution to be a paradigm case of group agency, and other instances of collective action will count as group agency insofar as they are sufficiently similar to that and other paradigm cases. In this section, we sharpen these pre-theoretical categories and set the stage for understanding their distinctive phenomenologies.

We begin with group agents. Paradigm examples of group agents are educational bureaucracies, large non-profit organizations, nation states, and perhaps even some mobs, crowds, and riots. We often say things like "the United States is pursuing its national interests," or "the University intends to implement a new policy." According to a straightforward reading of these statements, we treat groups of human beings as if they are singular agents, which as a

---

[1] List and Pettit (2011) are careful to note that there are many types of group agency. These can be mapped on a spectrum. Thus, group agency should be understood to be a threshold concept with the different agential types having uneven distribution and fuzzy boundaries. Similarly, joint agency is often identified following Gilbert's (1989, 1990) early work on plural subjecthood, with actions that can only be fully rationally explained by positing a plural subject. However, this covers a wide array of activities from spontaneously deciding to walk together with a friend, to pre-planned skilled performances of music or dance. So, "joint activity" should also be understood on a spectrum with fuzzy boundaries and many key thresholds. And, of course, the two categories can overlap, as joint activity can be subsumed under a group agent's activity just as individual activity can be.

group have intentions, attitudes, beliefs, and so perform actions. This can be puzzling. How can a group be an agent over and above the individuals who comprise it?

The view that talk of group agency is either metaphorical or mistaken, sometimes called eliminativism, dominated much of 20th century social metaphysics. More recent theorists have pushed back against the eliminativist tradition, returning, sometimes explicitly, to an earlier tradition of strongly emergentist accounts, where the group agent has an autonomous force over and above the sum of its parts (for example, Bratman 1987, 2014; Coleman 1974; Gilbert 1989, 1990, 2001; Hager 1989; Tollefsen 2002a, 2002b, 2015). Recently, List and Pettit (2011) have argued for a middle ground view that rejects both eliminativism and strong emergentism. They understand agency in functionalist terms, where a group that acts on a single robustly rational collection of attitudes rightly counts as an agent (p. 75). They argue that group agents can be organized so that collections of individual attitudes are aggregated into group level attitudes. This moderate position means that group agents wholly supervene on the contributions of their members, and hence are ontologically dependent. Unlike the eliminativist extreme, however, List and Pettit's (2011) account allows for the group agent to exercise non-redundant autonomy. This autonomy is in a sense epistemic, but non-trivial.

On this view, if we only look at individual members then we cannot understand group dynamics, even in idealized cases, because the relationship between individual attitudes and group attitudes is complex. In fact, List and Pettit (2011) observe that sometimes the aggregation function that combines individual attitudes into group attitudes can generate a rational practical outcome that no individual agent prefers.

This influential account of group-agency provides the foundation for our discussion; however, we wish to caveat two of its features. First, List and Pettit (2011) assume a

functionalist understanding of agency, in which agents must have representational states, motivational states, and the ability to process their representational and motivational states to intervene in the environment whenever that environment fails to match a motivating specification. They call this the "core" idea of agency. We think this picture of agency is inadequate for characterizing most individual human activity. Second, List and Pettit's rejection of emergence is too hasty. They seem to reject emergentist accounts out of a desire to avoid a "mysterious" or metaphysically "extravagant" account of group agency. We think, however, that the notion of weak (or "scientific") emergence can explain how complex entities like group agents are wholly ontologically dependent, and yet they can nonetheless constrain their members.

Both of these caveats become important in sections 4 and 5. For now, however, we turn to comparing group agency with joint activity.

In contrast to group agency, joint action involves no separate group entity. Whereas group agents are defined by the structural relations between their members, including the aggregation functions that deliver group level propositional attitudes, joint activity involves first and foremost interpersonal relations between individuals. Paradigm cases of joint activity include partners in a dance, people preparing dinner together, or two authors collaborating on a single piece of written work. This kind of joint activity, defined primarily by symmetrical relations between individual agents, has been the main focus of recent analytic philosophical work on collective action. Theorists have been centrally concerned with shared attitudes and intentions, and the debate has often split over whether these can be reduced to the attitudes and intentions of individuals who comprise the group (see, for example, Searle 1990; Roth 2003; Bratman 2009, 2014). For now, we set these debates aside. We assume there are shared attitudes, and that these

shared attitudes are indeed necessary to explain how multiple agents can act together in a single activity. In fact, this is one of the key features of joint action that separates it conceptually from group agency: in joint action, agents participate, at least in part, in virtue of having some attitude or collection of attitudes with sufficiently similar content across participants to unify the collective activity and make it more than just a simultaneity of individual actions; group action, on the other hand, involves group level attitudes that may in principle be entirely different from the attitudes held by the individuals comprising the group. We think that joint activity and group activity are distinct kinds of collective action, and we suspect that they cannot be unified into a single theory. In any case, what matters for our current purposes is the phenomenological differences between agents acting within a group agent and those acting jointly with others.[2]

---

[2] Though it is helpful to see joint actions and group actions as two distinctive types of collective activity; it is equally important to remember that common group terms do not always neatly map onto these types. For example, the concept of a "team" can cover both instances of what we call joint activity as well as group agency. The shared evaluative outlooks of an intramural football team composed of friends and roommates might be a clear case of joint agency; whereas, a large NFL team apparatus often (though perhaps not always) functions more like a group agent. As we discuss later in the paper (and briefly mentioned in footnote 1), this is further complicated because joint agency can be *embedded* within a larger group agent. For example, the 1991-1992 Chicago Bulls clearly had a cohesive, jointly agential, team dynamic at the level of the players *while also* functioning as a corporate entity exercising its group agency with brand deals and team press releases. This doesn't even begin to engage with the notions of "teams" employed in other domains of human activity (i.e. medical teams, military teams, business teams, etc.) nor does it consider non-ideal cases where the team develops a shared evaluative outlook but through coercive "team building" practices that could also alienate individual members. As with many instances of social metaphysics, the details matter when it comes to how these activity types are applied to concrete cases of human action; however, this does not undermine the usefulness of the categories themselves. We thank an anonymous reviewer for drawing our attention to these considerations.

We start our exploration of these phenomenologies by identifying two distinct ways in which any collective action can be experienced. As with the distinction between group and joint action, we take this difference to be pre-theoretical. On the one hand, an agent might experience collective action as an *augmentation* of her own agency. That is, she might experience collective action as an expansion of her agential powers, both in terms of expanded outcomes she is able to achieve and in terms of the array of worthwhile activities she is able to pursue. Moreover, an agent often, through collective action, comes to see more clearly the point of those very activities. On the other hand, an agent might experience collective action as a *diminution* of her individual agency. In this kind of experience, the ways that collective activity restricts the individual's freedom become especially salient. In collective activity, the agent can come to feel disconnected from her actions, and she may lose sight of the point of her activities. While it is possible that experiences of agential diminution could be positive, as when an agent feels a sense of belonging to something larger than herself, we are interested in cases where such experiences are negative and indicate problematic or non-ideal features of the collective activity that generates them.

This distinction between experiences of agential augmentation and diminution calls attention to the normative elements of agency. List and Pettit (2011) evaluate group agents both functionally and ethically. They call group agents that violate ethical norms "degenerate" (for example, a dictatorship) whereas "healthy" group agents have protections to support the rights of their members (for example, a liberal democratic state). We follow List and Pettit in considering the normative dimensions of collective action, but we exclude certain limit cases from our discussion. A dictatorship, for example, pushes the boundaries of what counts as a group agent. Similarly, we ignore certain edge cases of joint agency. A social bully who makes all the

decisions and indoctrinates a minion into adopting the bully's own evaluative judgments might present a kind of collective activity. However, the evaluative attitudes in such a scenario seem not to be shared in the right way. Such activity, because of the oppressive and determining power of the bully, pushes the boundary of our pre-theoretic notion of joint action.[3]

Keeping these exceptions in mind, we focus on non-ideal group agents because, although either kind of agential experience—augmentation or diminution—can occur in any case of collective activity, we think that experiences of diminution are especially common among members of complex group agents. We suspect that most group agents are not optimally configured, and so most of people's phenomenal experiences are of non-ideal groups, making the experience of diminution more likely. Our discussion of joint agency focuses on healthier cases, in part because we think that most actual experiences of agential augmentation occur in the context of joint agency, and in part because, in contrast with group agency, degenerate instances of joint agency seem further from the paradigm cases that define joint activity.

3. The Phenomenology of Group Agency

Consider the following case:

> Kira is the assistant head of the office of Diversity and Equity at a flagship state
> university. The goal of her office is to handle bias reports, provide guidance to
> academic units regarding departmental climate and job searches, provide diversity
> training, and ensure compliance with state and federal laws pertaining to diversity

---

[3] Thanks to an anonymous reviewer for drawing our attention to these problem cases.

and educational accessibility. Kira believes strongly in the value of diversity for academic institutions and cares deeply about the welfare of students. Recently, racist messages were found graffitied on the side of an important university monument; moreover, a conservative group of legislators within the state has taken umbrage with supposed "unfair diversity policies that infringe on free speech" and is attempting to cut funding for the diversity office. In response to these events, the University makes statements about its "position on diversity issues" and proposes some actions (including minor cuts to appease the legislators and a "Day of Diversity" event to show solidarity with students impacted by the racist message). Kira feels as though the university's pronouncements are somewhat disconnected from the complexity of the problems. Moreover, some of the responses seem insensitive to the actual values of diversity and equity that she holds and instead seem aimed at balancing student welfare with maintaining legal and financial stability. Despite her administrative position, Kira's day-to-day work experience seems only distantly related to her values. She mostly answers emails, attends meetings, and reads bias reports; often, she merely points people to resources already available on the University website. Even worse, the listening sessions with academic units frequently involve burden shifting, where an academic department asks for "help" from the office of diversity, but really wants the necessary work of diversifying to be done by someone other than themselves. Kira has trouble seeing how her day-to-day work really contributes toward the goals that convinced her to take up that work in the first place. The crisis has clarified how her work can feel like a kind of "maze" that she has to navigate in

order to "get to what matters." She often feels powerless to push ideas beyond
meetings into action.[4]

Before we analyze this experience of agency, we need to clarify some
terminology. First, the term "sense of agency" is intended to capture the feeling of being
the source of one's actions (Pacherie, 2008). Gallagher (2000, 2012) distinguishes what
he calls "sense of agency"—the sense of having *caused* an action to occur—from "sense
of ownership"—the sense of being the one who experiences the action. We think that
both of these elements are important features of a complete sense of agency.[5] Second, an
evaluative judgment is a representation of value. The total of an agent's evaluative
judgments, along with her dispositions to make such judgments, constitute her *evaluative
outlook*. Importantly, the dispositions to form evaluative judgments that make up an
evaluative outlook are not brute, but are the product of an *evaluative understanding* (we
will have more to say about understanding shortly).

Now, the experience described in the case above exemplifies feelings of diminution of
agency, as we described it in the previous section; more specifically, this is an experience of
what we shall call *alienation*: the feeling of being disconnected from one's activity.[6] Alienation
occurs when one's actions feel foreign to oneself. We hypothesize that feelings of alienation

---

[4] This case is an adapted conglomeration of several real-world cases conveyed to one of the authors by several close
friends. Names and details have been changed for privacy.

[5] The idea of treating the sense of agency as a complex concept can be traced to Bayne and Levy (2006) as well.

[6] We have borrowed and adapted this term from Marx. Though the concept appears in various places, the *locus
classicus* is Marx (1844/1988). For a more extensive philosophical discussion of what it is to identify or be
wholehearted in acting see Frankfurt (1988) and Watson (2004).

come from dissonance between one's ongoing activities and the evaluative judgments that guide (or ought to guide) those activities. In ideal rational action, we contend, one's activity is directly guided by judgments of value. Alienation is the phenomenal component of a mismatch between those evaluative judgments and the actions that they function to guide. One's evaluative outlook is a central constituent of one's practical identity. Thus, if the activities into which an agent invests her time and energy routinely fail to match the evaluative judgments whose function it is to guide such activity, it should be no surprise that there is a characteristic phenomenal quality to this experience, which is closely related to one's sense of agency.[7]

The case above illustrates how this separation can occur within the context of a complex group agent. Kira judges that the goals of the organization she works for are good to pursue, but her day-to-day activities are far removed from those goals, and do not seem to be (directly) guided by them. Her actions have only a distant connection with the evaluative outlook that informs them. This discontinuity occurs in part because when a group agent acts, the individual agents whose actions collectively compose that group action may not experience the group agent's action as their own. For example, Kira signs a memorandum that effectively institutes a

---

[7] For an account of alienation that has some similarities with our view see Schroeder and Arpaly's (1999) "Alienation and Externality." Their discussion, unlike ours, is centered around Frankfurt's (1971) desire-based model of the will; even so, they argue that the experience of alienation is connected to one's self-image or self-conception. In their view, the externality of certain desires to a person is *explained* by experiences of alienation wherein desires are felt as incongruent with the person's self-image. Our view shares with theirs a recognition that experiences of agential alienation are intimately related to the sense of practical identity; however, we differ from Schroeder and Arpaly in that we think (i) agential alienation is *specifically* related to a mismatch in evaluative judgements, and (ii) evaluative outlooks are partly *constitutive* of practical identity. We thank an anonymous reviewer for drawing our attention to this helpful overlap.

new university policy; she understands that she helped cause that new policy to go into effect, but she does not have a sense that *she* adopted the new policy. The university performed an action distinct from any of Kira's actions.

Feeling a lack of ownership of the actions performed from within a group agent contributes to the experience of alienation. The individual agents do not experience the group agent's actions as their own. Thus, they may not feel the appropriate connection between their individual actions and the evaluative outlook that informs those actions. Without experiencing that connection, the individual's actions feel disconnected from her practical sense of self, since her evaluative outlook partly constitutes that sense of self. Thus, she experiences alienation.

We propose that this feeling of alienation is common within paradigmatic complex group agents. With this notion in hand, we can begin to see more clearly how joint agency involves a different phenomenology. But first, we will consider in more detail why group agents cause their members to experience alienation.

4. Humean Agency, Weak Emergence, and Degenerate Group Agents

We argue that two features of paradigmatic group agents incline them to result in agential alienation: (i.) their Humean evaluative attitudes and (ii.) their weakly emergent structure.

(i.) Humean Agency, Aggregation Functions, and Evaluative Judgements

What we call "Humean" theories of agency have three key elements. First, for any behavior to be an instance of agency on a view of this kind, it must be guided by a

representation. This guiding representation is a mental state that captures some state of affairs, but not a state of affairs that actually obtains. Rather, it is the state of affairs to be brought into being by the agent's action. So on this view, agents must be able to represent the world as they want it to be.

Second, for Humean theories of agency, whatever actually moves the agent to act cannot be part of the representational content of the mental state. Representational content, the thinking goes, merely depicts ways the world could be, and there is nothing within a mere description of a possible state of affairs that explains why that state of affairs should be preferable to any other. In other words, an agent's representation of the state of affairs that she is attempting to bring about, on the Humean model, must be purely descriptive; it cannot have any evaluative content at all.

Third, and finally, Humean views claim that the content of the representational mental states that move us to act must be exhaustively constituted by the proposition towards which one has an attitude. Any desire can be expressed fully as a desire *that* some state of affairs obtain, without any loss of information. Once again, the representational content of a desire is entirely non-evaluative.

Contrast this view with what we shall call an "Aristotelian" model of agency.[8] On an Aristotelian view, some representational mental states are *apprehensions of value*. These

---

[8] We intend the labels "Humean" and "Aristotelian" to capture two general families of views. There are, however, undoubtedly philosophers who consider themselves Humeans who do not share all of the commitments we lay out, and the same is true of Aristotelian views. Moreover, some philosophers hold views that do not fall clearly into either category. For example, Bratman (1987) explicitly rejects belief-desire psychology as a basis for explaining intentional rational action, and so his account may not be well captured by the commitments that we term Humean; however, he presents a functionalist account of agency, which is clearly not in the family of views we call

evaluative representations have evaluative content that makes them necessary for rational explanations of action. This means that in order to make sense of an action, we have to be able to understand *what the agent sees as good*. But importantly, one need not have any particular propositionally specifiable state of affairs in mind in order to apprehend the value of some activity and thus be moved to act. An apprehension of value can move an agent to act even if she lacks a clear view of what state of affairs will obtain once she has completed the action. Indeed, she need not even have a clear view of what constitutes "completing" the action.

As Brewer (2009) argues, many human practices involve engaging in activities the value of which we cannot fully understand at the start. Realizing this value does not involve making the world conform to a preconceived state of affairs; rather, it involves engaging diachronically in the activity itself. This means that the agent's actions over time are guided by a continually developing understanding of the value of the activity, and this value cannot be conceptually separated from the activity. On this view, having a desire means seeing some activity or object *as good*, which means representing the object of the desire in a way that has fine-grained and ineliminable evaluative content. The specific nature of this evaluative content, and the agent's understanding of it, guide the agent's actions. Those actions then bring the agent into new relations with the value being perceived, thus deepening her understanding of that value and drawing her toward further actions. Brewer argues for this view, in large part, by arguing that without the notion of a constantly developing understanding of what is independently valuable, we cannot make sense of the unity of our activities, and indeed of our entire lives. Humean views

---

Aristotelian. We still use these labels because (a) Hume (at least as he is often interpreted) famously denies that representations can have evaluative or motivational content, and (b) Talbot Brewer, whose work we rely on for developing an alternative view, presents his position as inspired by Aristotle's ethics.

leave us with an understanding of human activity that is fragmented into atomized individual actions held together, if at all, only by arbitrary preferences or brute dispositions.[9] Functional and Humean conceptions of agency, because they do not have the resources to explain the role of evaluative understanding in rational agency, cannot capture important aspects of human activity.[10] As such, though we agree with List and Pettit that *group* agency is functional and Humean, we think that this is not a complete account of *individual* human agency.

Now, many questions remain unanswered about how to flesh out such an Aristotelian model of agency, and we do not claim to present anything like a full-fledged theory. Recall, however, our stated method of arguing from explanatory power, interest, and fruitfulness. Adopting an Aristotelian model of individual agency enables an interesting and fruitful explanation both of the phenomenology of joint action, as well as of how that phenomenology differs from the phenomenology of group action, and that itself speaks in favor of the view.

Before moving on, we should say something about the notion of *understanding* at work here. Whatever substantive theory of understanding one endorses, we think it is clear that

---

[9] See Brewer (2009), especially chapter 2.

[10] There are sophisticated neo-Humean theories of action that attempt to circumvent such difficulties. For example, Schroeder's (2007) *Slaves of the Passions* defends both a Humean theory of reasons as well as reductive metanormative realism. Unfortunately, fully engaging with such views individually is outside the scope of this paper. We briefly note that because of Schroeder's commitment to metanormative reduction any evaluative understanding is ultimately explained by non-evaluative relations and psychological states; however, our general claim is that this strategy merely *explains away* the target of analysis. Schroeder (2012) elsewhere claims that ultimately he is appealing to a *parity* argument; his view does not require the additional evaluative baggage of other more normatively laden views. We agree, but worry that in making the view more parsimonious the ultimate target of analysis has been lost; sometimes explanatory fruitfulness and adequacy outweigh parsimony.

understanding is a feature of *thought*. Thoughts are mental states, and although thoughts have propositional structure, they are not identical with propositional attitudes. Thus, even though it may be correct to attribute propositional attitudes to group agents, that does not mean that the group has thoughts. Even though a group agent might have various propositional attitudes, and might acquire those attitudes according to coherent patterns and principles, it would not be correct to say that the group agent understands anything, at least in the sense that we mean. Indeed, it is precisely because understanding is a feature of thought that it is fit to play the role that it does in explaining our engagement in dialectical activities: engagement in dialectical activities involves *continuous practical thought* (or at the very least a capacity for continuous practical thought), not just the formation of motivationally efficacious propositional attitudes.

Finally, we want to be careful not to overintellectualize understanding. As Brewer (2009) presents it, our evaluative understanding shapes what things appear to us as good, often without any deliberate reflection, just as our various theoretical understandings shape how things appear to us in general. This suggests that understandings can be, and often are, implicit. One need not be able to explain or articulate clearly one's understanding in order for that understanding to shape how one perceives the world.

Now let us return to Humean agency. We contend that group agents are genuinely Humean; that is, group action is explainable primarily in terms of propositions about conceptually independent states of affairs that the action aims to bring about.[11] For example, the university can *produce* a statement that proclaims its commitment to students harmed by discriminatory acts; but the university cannot *participate* in the ongoing activities of empathizing with, respecting, or loving students. We hypothesize that group agents cannot be guided by

---

[11] That is, the states of affairs to be brought about are conceptually independent from the activity itself.

complex apprehensions of value because to be guided by an apprehension of the value of an activity is to be in the process of developing an ongoing understanding of that activity's value. One is drawn initially into action by an inchoate sense of the value to be realized, and only by engaging in the activity can one further one's understanding of that value. Brewer (2009) calls activities of this kind *dialectical*.[12] But group agents (or at least, group agents of sufficient size and complexity; there likely are borderline cases) cannot have ongoing and constantly developing understandings of the value of what they are doing. This is for two reasons, one structural and one metaphysical.

First, consider how List and Pettit (2011) describe the "core" idea of agency: (i.) agents have representational states that depict how things are in the environment; (ii.) agents have motivational states that specify how they require things to be in the environment; and (iii.) agents have the capacity to process their representational and motivational states, leading them to intervene suitably in the environment whenever that environment fails to match a motivating specification. On this view, a representational state is an intentional attitude whose object is a proposition concerning a state of affairs depicted as true, whereas a motivational state is an intentional attitude whose object is a proposition concerning a *possible* state of affairs that the agent *wants* to make true (p. 21). For these attitudes to amount to agency, on this view, they must also be held to certain basic standards of rationality. But once these criteria are met, we can generally describe representational states as "beliefs" and motivational states as "desires."

---

[12] The term "apprehension of value" comes from Brewer (2009). Elsewhere in this essay we use the term "evaluative judgment" to refer to representational mental states, typically beliefs or belief-like states, that take these apprehensions of value as part of their content.

Group agents, then, on List and Pettit's view are counted as agents just in case they possess those three "core" capabilities. These capacities come to be instantiated through an "aggregation function," which merges the intentional attitudes of the individuals into what the group agent believes or desires. These functions are specified by the organizational structure and take many forms (e.g. majoritarian voting, distributed premised-based deliberation, straw-polling, etc.). Finally, List and Pettit (2011) assume that "whatever the group's organizational structure is, it must lead the group to form binary attitudes, not attitudes that come in degrees of strength" (p. 37). This means that a group agent's beliefs are judgements that either p or not-p is true and a group agent's desires are preferences that either p or not-p be true. This is because, in general, for an aggregation function to be successful it must be able to coordinate the members of the group agent as they form and enact proposed group attitudes. This communicability requirement will, List and Pettit claim, lead a group agent to binary attitudes because of their clarity. An on-off judgment about whether the group believes something is the case or an on-off preference for whether the group desires a certain outcome is easily communicated and enacted within the confines of a group's organizational structure.

We can now clearly see why group agents of sufficient size and complexity *must* be Humean agents. The core idea of agency allows for the aggregation function to be successful because it presents the objects of action as *propositions* regarding possible states of affairs that are "to be produced." Moreover, these group attitudes must be binary attitudes in order for the attitudes of members to be properly aggregated. If organizational structures like voting are to work, the courses of action under-consideration must be *well-defined*, which means they take a propositional and binary form. These features make group agents inapt for an Aristotelian interpretation. They do not have the requisite nuance to engage in dialectical activities, where the

object of action is *not* a non-evaluative proposition concerning a possible state of affairs. This is the "structural" reason why group agents are Humean; but there is a metaphysical reason as well.

Group agents are purely functional entities. There is nothing *that it is like* to be a group agent, no conscious mind to apprehend value in any manner *other than* syntactically as a propositionally mediated representation.[13] To be sure, evaluative variability might be communicated to the constituting members via deliberative feedback mechanisms within the organizational structure (e.g. a committee meeting about qualitative surveys measuring the public's responses to their recent actions, etc.). These individual members might then have nuanced engagement with apprehensions of value. But crucially once their newly formed attitudes are fed into the aggregation function of the organizational structure they must be "propositionalized" (that is, fully specified and made binary) to allow for group-wide reasoning. Moreover, such a fully specified proposition cannot have any evaluative content of the kind required by the Aristotelian model of agency because such evaluative content is necessarily mediated by an understanding of the good, which is continually under development through the agent's ongoing activities. This kind of understanding is a property of *thoughts*, which are mental states whose structure and content cannot be fully captured as binary propositional attitudes.[14]

---

[13] To be clear, though we present this as the most widely accepted position, there are reasons to investigate the possibility, however remote, of not only group agency but group *experiences.* For an interesting discussion, see Pacherie (2017).

[14] This is true even if thoughts always have propositional structure. Perhaps a thought is always partially constituted by a propositional attitude; that does not mean that having the propositional attitude is the same as having the thought. Moreover, it is not clear that the propositional content of thoughts is always, or even usually, fully specified and binary in the way that group-wide reasoning requires.

Since group agents (as far as we know) lack unified minds with actual thoughts, they cannot have this kind of ongoing evaluative understanding.

This explains why the alienation we identified in paradigmatic experiences of group agency is felt as a lack of a sense of *ownership*. For example, Kira recognizes that the actions she individually performs as a constituting and active member of the university are indeed *hers*. As Pacherie (2014) reminds us, most cognitive models of the sense of agency rely on a congruity principle, which claims that a felt sense of agency occurs when there is a match between two cues in our agential architecture (either at the pre-reflective level of our motor processes or at a more complex reflective level). In Kira's case, this full sense of agency occurs locally (e.g. she initiated sending the email regarding draft-language for the university's response to racism, and her expectations regarding her performance of typing that email and sending it were met). Even at the group level she played a causal role in producing the university's final statement, as it contains some of her draft language. The university, however, cannot share in Kira's evaluative understanding. Group agents, because of their Humean structure, act to bring about binary states of affairs that are described, for the purposes of group action, entirely non-evaluatively. The only "evaluative" element in the workings of the group agent is the arbitrary aggregated preference that the state of affairs in question be brought about.

Again, evaluative outlooks are constitutive of our practical identity. That our actions are congruent with what we judge to be good is part of what makes our activity intelligible as the doings of an integrated agent rather than merely a fragmented set of happenings. But the university is not the right sort of thing to have an evaluative outlook, since it is *necessarily*

Humean in its structure and has no conscious mind.[15] This incongruity is seen in how Kira feels

that the "pronouncements are somewhat disconnected from the complexity of the problems" and

that "some of the responses seem insensitive to the actual values of diversity and equity that she

holds." Kira has a nuanced and dialectical understanding of the value of diversity, whereas the

university must (in virtue of its Humean structure) cumbersomely produce particular states of

affairs that are able to be *clearly* represented on the front end.

<div align="center">(ii.) Weak Emergence and Agential Freedom</div>

An additional disconnect helps further explain the alienation that occurs within complex

group agents; this time it is grounded in how group agents are, plausibly, weakly emergent

entities.

List and Pettit (2011) explain the metaphysical relationship between group agents and

their members in terms of supervenience. But supervenience as normally specified is too weak a

relation to do much explanatory work. The standard view is that A supervenes on B if there is no

A-change without a B-change. Though this does clearly capture some sort of relation, it does not

specify *why* this interaction obtains (see Horgan 1993, and Kim 1997, for two early criticisms

along these lines). In particular, one wants to know what direction the arrow of dependence runs.

---

[15] There may be some weaker notion of evaluative outlook that could apply to group agents. If, for example, one

thinks of an evaluative outlook merely as a collection of propositional attitudes and dispositions to form such

attitudes, then a group agent could have an evaluative outlook in this weaker sense. But in the sense we mean,

having an evaluative outlook implies having an evaluative understanding, and understanding implies thought. A

group, as far as we know, does not have a mind, and thus cannot have thoughts.

We argue that an appropriately restricted understanding of weak emergence enables a fuller explanation of how group agents both depend on and constrain the powers of their members.

Before continuing, it is important to note that the concept of "weak emergence" originated in the context of metaphysics and philosophy of science; however, we are applying the concept at the social level. Where our use of the concept differs from its application to the physical sciences or philosophy of mind we will make note.

An emergent entity is (in some sense) *dependent* on the base entities that constitute or support it, while at the same time being (in some sense) *autonomous* from those entities (Corradini & O'Connor 2010, p. 3). We claim that weak emergence, a version of Wilson's (2015) "proper subset" approach, somewhat adjusted to make it more amenable to social metaphysics, is the best way to characterize the relationship between a group agent and its members. Wilson argues that the strong/weak distinction tracks whether the emergent property has at least one more power than its base, or if it has fewer powers than its base (p. 362). She specifies the notion as follows:

> *Weak emergence*: Token higher-level feature *S* is weakly metaphysically emergent from token lower-level feature *P* on a given occasion just in case (i) *S* synchronically depends on *P* on that occasion; and (ii) *S* has a non-empty proper subset of the token powers had by *P*, on that occasion. (p. 362)

Wilson means "powers" to be understood in as neutral a manner as possible, since for her talk of "powers" is simply shorthand for what causal contributions a feature makes to achieve an effect

(p. 354). This is meant to track the uncontroversial thesis that the way an entity is contributes to what the entity can do.

This proper subset strategy was first posited and developed by Wilson (2011, 1999) to solve puzzles regarding mental causation; however, it is useful as a general characterization of weak emergence. It is motivated, primarily, by cases of multiple realizability. The idea is that multiply realized powers are those in the intersection of the sets of powers of the emergent entity's realizing constituents, and hence these powers are part of a proper subset at the token level (p. 358).

For example, suppose that some attitude of a group agent, call it, *AG*, is multiply realizable. Further, suppose *AG* is the university's disjunctive desire to support diversity and ameliorate public tension, which causes an action, *A*, of making a media statement regarding the importance of both diversity and free speech. Here, *AG* is realized by the efforts of some set of the university's constituting persons *P*. But suppose it had instead been realized by some other set of constituting persons *P**. Would *A* still have occurred? It seems so because the differences between *P* and *P** do not matter for *A*; all that matters is *AG's distinctive power profile*, which contains the powers crucial for *A*. To make this concrete, imagine two possible scenarios: in the first scenario; Clayton, a communication director in the university, is tasked with releasing a statement about a new recruiting initiative in the school of engineering. Clayton does not really care about matters of student recruitment in engineering, but he dutifully follows procedure and releases a well-crafted statement to local and state media regarding the university's intentions. In the second scenario, Charlene, a communication director in the university, is tasked with releasing a statement about a new recruiting initiative in the school of engineering. Charlene cares deeply about matters of student recruitment in engineering and she enthusiastically follows

procedure and releases a well-crafted statement to local and state media regarding the university's intentions. Even though there are important differences between the people constituting the group agent's attitude (*AG)* in these two scenarios, this does not change the university's action. Instead, the *structures* of the university (its "procedures" and "offices") select for the powers necessary to realize its actions regardless of whether those powers are found in a person like Clayton (*P*) or Charlene (*P**) so that the university still performs the same action, *A*, regardless. This sort of example gives us, Wilson (2015) argues, a principled reason for taking *AG* to be efficacious with respect to *A* in a way distinct from *P* or *P** (p. 361).

The crucial insight here is that such an entity is weakly emergent when it exhibits *fewer* powers, or to use Wilson's terminology, fewer "degrees of freedom" than its base parts. Of course, these theories were developed in the context of philosophy of mind, so the precise definition of "degrees of freedom" in physics is inapt for social metaphysics; however, the thought that a weakly emergent entity comes into being by *structuring* its constituent parts, so that their powers are restricted in ways that allow for the emergent entity to manifest distinctive higher level powers is useful.[16] Consider the wide variety of activities that a human person can do. As rational agents, our scope of action is wide for each individual; however, when *constituting* a group agent there are usually norms that *constrain* what someone may legitimately do in that context. Kira cannot just release her own personal statement about diversity on behalf

---

[16] We emphasize that it is the structural similarities between the concept of a weakly emergent social entity and a weakly emergent physical entity that make the comparison apt. However, there are differences in the exact application of this structure. There may be a kind of conditional necessity, which holds that *if* persons are embedded within some social structure *then* a weakly emergent group agent comes into existence. But, unlike fundamental instances of physical emergence, this lawlike emergence at the social level always assumes that all other things are equal, and often they are not equal.

of the university without going through the proper channels. (If she did, it simply would not be an action done by the university). Their actions are limited by the structure of the emergent entity they are constituting, and through these constraints the group agent leverages their powers towards *its* distinctive ends. So, by limiting Kira's expressions merely to sending an email with a draft version of the diversity statement through the proper channels, the university gains the power of making a statement *itself*.[17]

This further explains the distinctive phenomenology of group agency. Kira's actions are disconnected from the actions of the group agent in part by the lack of control she has over *how* she performs her daily tasks. This lack of control is grounded in the weakly emergent structures, which constrain the agency of individuals in order to realize the group agent. Moreover, the individual is replaceable because the organizational structures that harness her contributions can be realized by a wide variety of persons as long as they have the proper subset of powers. Kira's contributions to the group agents' actions are not distinctively and wholly *her* agency, but instead merely a part of her agential powers; hence, they are replaceable by anyone who's agential powers can fill the same structural need.

A natural question concerns whether the kinds of constraint that weakly emergent entities impose on their constitutive members distinctively explains the sort of alienating constraints imposed by group agents. After all, joint activity also imposes constraints. Suppose I am running

---

[17] Kira *could* eschew university structure and release her own statement about diversity, perhaps even appropriating trappings from the university to make it seem more "official." After all, the group agent merely structures her agency; it does not determine it. But social pressures, spoken and unspoken norms, perhaps even disciplinary oversight, *push* human agents down certain routes of action, encouraging only the proper-subset of their agential powers that are necessary for the group agent to function.

with a friend; it would not be "running together" if I run ahead and leave my friend behind. I am constrained, insofar as I am the plural subject of a joint activity, to match my pace with my companion. Given this, what distinctive explanatory work is the concept of weak emergence doing?[18] Crucially, it is not the mere *fact* that there are constraints that needs explanation, but instead *how* these constraints might be felt differently depending on whether one is participating in joint or group activity. We believe that weak emergence can provide this explanation because of where it locates the source of the constraints.

This is best seen by way of contrast. In the case of joint activity, like "running together," the source of constraint is my own agency guided by the evaluative understanding that I both affirm and share, insofar as possible, with the other constituting members of our joint action. I am neither "fragmented" from my agential powers, since my evaluative understanding of the joint activity is always under development; nor am I alienated from the ends of my activity since the constraints placed on me by these ends have their source in my own agency. In contrast, with weakly emergent group agents, the constraints imposed have their source in the emergent entity's structure. Therefore, group agency will likely be experienced as a fragmenting of a constituting individual's powers, since only a subset of those agential powers are relevant to the group agent. It will also likely be experienced as alienation since the ends are not selected according to the individual's own evaluative outlook but instead by the propositionally specified states of affairs that the group agent selects for.[19]

---

[18] Thanks to an anonymous reviewer for raising this concern.

[19] Of course, the group agent's ends and an individual's own evaluative outlook may be *related* in various ways, but this is not the same as sharing an apprehension of value.

These two features of paradigmatic group agents—(i.) Humean agency and (ii.) weakly emergent structures—are the underlying explanation for *why* the persons who comprise these group agents experience alienation, a sense of *diminishment* of their own agential powers. With this phenomenon now clearly defined we can use it as a contrast case for joint agency. In the following section we highlight how the experience of joint agency differs from the experience of constituting a group agent.

5. What it is Like to Act Jointly: Joint and Group Agency Compared

Joint agency, we think, involves shared apprehensions of value among the individuals acting together. Just as the phenomenology of group agency is partly explained by the disconnect between the evaluative outlooks of the individuals and the mechanisms that guide the group agent, the different phenomenology of joint agency suggests a close connection between the evaluative outlooks of participating individuals through their shared evaluative judgements and understandings. Joint agency typically does not create experiences of alienation. Furthermore, we suggest that experiences of successful joint agency are especially fulfilling because they are experienced as *enhancements*, rather than diminutions, of individual agency.[20] We hypothesize

---

[20] Could a group agent also contribute to experiences of *agential enhancement*? We cannot explore this question in detail, but we briefly mention two considerations: First, as we discuss in the conclusion, the healthiest group agents enable, support, and coordinate instances of joint agency within their structure in order to constitute their ultimate actions. Insofar as a group agent is *necessary* for enabling and coordinating these joint agencies we may rightfully say that that group agent contributes to an experience of agential *enhancement*, though it is only *through* the instances of joint agency. Second, suppose that the outcome of a group agent's action is an outcome also desired by one of its constituting individuals. Would this be experienced as an *enhancement* of individual agency? We think the

that this is because (i) joint agency does not involve an emergent entity as the subject of the final action, and (ii) jointly acting agents (at least in successful paradigm cases) share an evaluative understanding that guides their activity. The absence of an emergent group agent explains why joint activity does not typically involve feeling a lack of ownership of one's actions. The shared evaluative understanding explains why joint action is often experienced as an enhancement of one's individual agency, and also what makes joint action different from merely synchronous individual activity.

To illustrate, we present another hypothetical case:

On a Friday afternoon, Kira and her friend Emillio meet in a local park to go for a long run after a stressful work week. Both enjoy regular runs, occasional races, and discussions about running. As they run together their pace "locks-in" to where they are running side-by-side in a similar rhythm. Individually, Kira tends to run a bit faster and Emillio a bit slower, but when running together they meet somewhere in the middle. If one of them starts to seem tired, the other is attentive to this and will adjust pace accordingly or perhaps give some verbal encouragement. As they continue to run, they each attend to their technique and pay attention to their companion's as well. But, as they continue they become

---

answer is probably no because the group agent is a separate agent. Even if its actions lead to outcomes which a group member also desires, the agential process is going to be alien to an individual's agency. This is contrasted with joint agency (in the ideal case) when two agents are *wholly* brought into a mutual evaluative understanding such that they both experience their agency as enhanced and extended. Again, in contrast, group agents only "use" the subset of an agent's powers necessary to constitute themselves, and because group agents lack thought they metaphysically *cannot* have a shared evaluative understanding with their group members.

engrossed in the activity, occasionally chatting about different routes they could

take. Sometimes, deciding to increase the pace as a fun temporary challenge, they

speed ahead, weaving seamlessly past other slower pedestrians. After an hour and

a half they reach the end of their circuit. They stop and cool down with some

stretching. They start to talk excitedly about one of the different routes they took

with a challenging hill they had to overcome; each begins offering suggestions

about how they could surmount steep hills more quickly and easily. Another 30

minutes passes as they just continue to discuss their run; finally Emillio laughs

and says, "It's so good to be able to run with someone who really 'gets it,' I feel

more excited about running after we hangout!" Kira concurs.[21]

This case has several important features. First, it is worth noting how attention functions in the

case.[22] Jointly acting individuals attend to not only their own behavior but also the relation of

their behavior to the actions of others. It is this "me-as-related-to-them" and its converse that

---

[21] If the reader does not find this (fictional) case compelling, it is easy to think of other activities with the same

dialectical structure: philosophizing, artistic projects, small business ventures, book clubs, and so on. Indeed, we

contend that this sort of activity makes up the bulk of deliberate human pursuits.

[22] Recent philosophical work on joint agency has been concerned with metaphysical questions about agency and

mind (see for example, Bratman 2014, Eilan 2005, and Heal 2005). One branch of that research program has

focused specifically on joint attention, and how to understand the epistemological mechanisms of attending jointly

to the same object with another person. Eilan (2005) points out that many theorists in philosophy and psychology

have found attractive the idea that "there is something utterly simple and basic about the transparency of our minds

to each other in the case of joint attention," but explaining that insight in a satisfactory way has proven very

difficult. In any case, the transparency of other minds does seem to be part of the phenomenology of joint action.

partly distinguishes joint activity from actions that "just so happen" to be occurring at the same time. Of course, if while running solo Kira happens upon another runner she still may pay attention to how her behavior relates to the other person's (perhaps she will slow down so that she is not awkwardly running beside them), but crucially in joint action the "me-as-related-to-them" is attended to because it is *constitutive* of the action being performed. This phenomenological difference, where I attend to others' behavior because I experience it as partially constitutive of *my own* action, requires explanation. It is here that the other feature of the case, the mutual *engrossment* in the activity's aim, becomes important. Apprehensions of value structure joint actions by being mutually held. The mere sharing of an evaluative attitude is insufficient. The sharedness of the activity must in some way be part of the content of the apprehension of value guiding the activity. Again, the mutual apprehension of value is partially constitutive of what makes joint activities *joint* in the relevant sense. After all, the mere fact that individuals are physically together, performing the same activity for the same reasons seems insufficient for their actions to be *joint*. Consider Searle's (1990) example of a group of people all running to the nearest shelter to escape oncoming rain; this would not be joint activity.

Our view provides promising resources for explaining what differentiates incidental acting together and genuine joint action. On this view, joint activities are individuated by the apprehensions of value that guide them. Many of the values that jointly acting agents apprehend have, as constitutive features, *that* they be pursued together. Moreover, one cannot fully understand the value of the activity without pursuing it. Thus, in order to be guided by these kinds of values, one must pursue them socially, not just in tandem with others, but with mutually

self-conscious sharing of evaluative understandings.[23] For example, deliberate activities of friendship simply cannot coherently be pursued accidentally in tandem. If they were, they would not be activities of friendship. It is the perceived value of the other person that guides, at least in part, the activities of friendship. This suggests that what separates genuine joint activity from accidentally synchronous activity is the content of the evaluative judgments that guide the activity, and the way that those evaluative judgments can have mutuality, and indeed the value of others, as part of their content.

Finally, we also claim that a shared evaluative understanding can explain how joint agency is paradigmatically experienced as an enhancement of one's individual agency. Recall that on the Aristotelian model of agency many of our activities are guided by apprehensions of value that necessarily start out inchoate. But one must have some beginning understanding of the value of the activity in order to be drawn to it in the first place. So the evaluative judgments that guide these dialectical activities are underspecified at first; they are incomplete and ongoing. We suggested earlier that the inability of complex group agents to act on such underspecified apprehensions of value says much about why those group agents can alienate individuals, but it also throws into relief how jointly acting individuals *do* share in and act on such inchoate and ongoing understandings of value. In fact, as Brewer (2009) makes clear, many of the central

---

[23] And since understanding, as we conceive it, is a feature of thought, sharing an understanding of an activity's value as joint activity implies that the individuals' thought contents include in some way the mutuality of the activity. It is an interesting question whether this shared thought content always requires metacognitive representation, where each agent represents the thoughts of the other in some way. If it does, then it would seem that fully fledged joint activity would not be possible for beings (perhaps very young children and most non-human animals, for instance) that lack a capacity for metacognition. Our account seems to suggest this restriction, though we do not intend to rule out the possibility of primitive versions of joint activity that do not involve metacognitive representation.

examples of dialectical activities depend for their nature on the fact that we engage in them together, and not just individually.

We think the fact that individual agents can share a single apprehension of value when engaging together in dialectical activities can help explain why those activities have the dialectical structure that they do.[24] It is precisely by sharing the same apprehensions of value that individual agents are able together to develop their understanding and appreciation of the value they are engaging with. Because acting jointly enables, or at the very least magnifies, our ability to develop ongoing understandings of value, it also expands our agential powers. A deeper understanding of the value within our activities allows us to engage with those activities, and probably other related activities as well, in ways that were not possible with lesser understandings. Acting jointly requires a diachronic deliberative process, wherein the agents clarify their apprehensions of value *for each other* mutually, intersubjectively, and dynamically. This, in turn, strengthens the congruity between their evaluative understandings and their actions, making their dipartite sense of agency—both *authorship* and *ownership*—stronger. This explains why joint agency is often experienced as especially meaningful. Moreover, unlike weakly emergent structures that impose *external* constraints that exploit only a specific subset of our agential powers for the sake of an end disconnected from us, these joint activities impose constraints on our agency that are *internal* to the evaluative understandings we ourselves hold and affirm.

There are, of course, complexities here. For example, cognitive scientists disagree about how and why the minds of other agents seem so transparent to us when we act and attend to

---

[24] See Brewer (2009), especially chapter 7.

things jointly (Eilan 2005). Nonetheless, we think this understanding of joint agency is promising and worthy of additional reflection.

6. Conclusion: How to Create a Healthier Group Agent

As we have shown, attention to the distinctive phenomenal character of group agency does much to highlight, via negative-relief, what is unique and desirable about joint agency. As we conclude, we briefly note that the theoretical usefulness of the dialectic between group agency and joint agency runs in the other direction as well. Our analysis of paradigmatic group agents painted a bleak picture of their capacity for supporting individual senses of agency; however, some experiences of group agency are certainly positive.[25] A group agent can be structured in such a way that it encourages many instances of joint agency among its constituting members.[26] A group agent that emerges *out of* an appropriately structured collection of joint agential projects is not a novel ideal; indeed, List and Pettit (2011) talk about this possibility explicitly. We think, however, that because of the essential mismatch between group agents and individual agency the benefits of an internal scaffold of joint agential projects as the constituting base of the group agent itself is vital for avoiding alienation within group agents of significant

---

[25] See footnote 1 above.

[26] Although many of these discussions are novel in the context of analytic philosophy, there are historical precedents that are often overlooked due to artificial disciplinary boundaries. In particular, we have found inspiration in the phenomenological and socio-political work of Edith Stein. For a useful introduction to her groundbreaking work on the phenomenology of community and the nature of joint action see Szanto and Moran (2015). We thank Isabelle Farineau for making us aware of these connections to Stein and how they might overlap with our thesis.

size and complexity. In the end, the phenomenology of joint agency suggests that *any* shared activity—from co-authoring a paper, to a university releasing a statement about diversity—should be structured as the kind of thing we the individual agents can identify with as belonging to us.[27]

References

Bayne, T., Levy, N. (2006) "The Feeling of Doing: Deconstructing the Phenomenology of Agency." In N. Sebanz and W. Prinz (eds.) *Disorders of Volition*. Cambridge, MA: MIT Press, 49–68.

Bratman, M. (1987). *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press.

Bratman, M. (1992). "Shared Cooperative Activity". *The Philosophical Review*. 101 (2): 327–341

Bratman, M. (1993). "Shared Intention". *Ethics*. 104 (1): 97–113.

Bratman, M (2009). "Shared Agency". In C. Mantzavinos (ed.) *Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice*. Cambridge University Press: 41–59.

Bratman, M. (2014). *Shared Agency: A Planning Theory of Acting Together*. Oxford University Press.

Brewer, T. (2009). *The Retrieval of Ethics*. New York: Oxford University Press.

Carpenter, M. (2009). "Just How Joint is Joint Agency in Infancy?". *Topics in Cognitive Science*. 1: 380–392.

Chaminade, T. Marchant, J.; Kilner, J. Frith, C. (2012). "An fMRI study of joint action–varying levels of cooperation correlates with activity in control networks". *Frontiers in Human Neuroscience*. 6: 179

Coleman, J. (1974). *Power and the Structure of Society*. New York: Norton

Corradini, A., O'Connor, T. (2010). "Part I: Introduction." In A. Corradini, T. O'Connor (eds) *Emergence in Science and Philosophy*. New York: Routledge. 3-7.

Dewey, J. A., Pacherie, E., & Knoblich, G. (2014). "The phenomenology of controlling a moving object with another person." *Cognition*, 132: 383–397.

Eilan, N. (2005). "Joint Attention, Communication, and Mind." In N. Elian, C. Hoerl, T. McCormack, J. Roessler (eds). *Joint Attention: Communication and Other Minds*. Oxford: Oxford University Press, 1–33.

Frankfurt, H. (1971). "Freedom of the Will and the Concept of a Person." *Journal of Philosophy*. 68: 5-20.

Frankfurt, H. (1988). *The Importance of What We Care About*. New York: Cambridge University Press.

Gallagher, S. (2000). "Philosophical Conceptions of the Self: Implications for Cognitive Science." *Trends in Cognitive Science*, 4 (1): 14–21.

Gallagher, S. (2012). "Multiple aspects in the sense of agency." *New Ideas in Psychology*.

Gilbert, M. (1989). *On Social Facts*. Princeton: Princeton University Press.

Gilbert, M. (1990). "Walking Together: A Paradigmatic Social Phenomenon." Midwest Studies in Philosophy, 1–14.

Gilbert, M. (2001). "Collective Preferences, Obligations, and Rational Choice." Economics and Philosophy, 109–120.

Hager, M. M. (1989). "Bodies Politic: The Progressive History of Organizational 'Real Entity' Theory". *University of Pittsburgh Law Review* 50: 575–654.

Heal, J. (2005). "Joint Attention and Understanding the Mind". In N. Eilan, C. Hoerl, T. McCormack, and J. Rosessler (eds), *Joint Attention: Communication and Other Minds*. New York: Oxford University Press, 34–44.

Heal, J. (2013). "Social Anti-Individualism, Co-Cognitivism, and Second Person Authority." *Mind*, 122 (486): 339–371.

Horgan, T. (1993). "From Supervenience to Superdupervenience: Meeting the Demands of a Material World." *Mind*. 102 (408): 555–586.

Kim, J. (1997). "The Mind-Body Problem: Taking Stock After Forty Years." *Philosophical Perspectives*, 11: 185–207.

List, C., & Pettit, P. (2011). *Group Agency: The Possibility, Design, and Status of Corporate Agents*. New York: Oxford University Press.

Marx, K. (1844/1988). "Estranged Labor." In M. Milligan (trans.) *Economic and Philosophic Manuscripts of 1844*. Amherst, NY: Prometheus Books: 69–84.

Meyer, M. van der Wel, R.P.R.D. Hunnius, S. (2016) "Planning my actions to accommodate yours: joint action development during early childhood." *Philosophical Transactions of the Royal Society B: Biological Sciences* 371: https://doi.org/10.1098/rstb.2015.0371

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, 107: 179–217.

Pacherie, E. (2012). "The phenomenology of joint action: Self-agency vs. joint-agency." In A. Seemann (Eds.), *Joint attention: New developments*. Cambridge, MA: MIT Press, 343–389.

Pacherie, E. (2014). "How does it feel to act together?" *Phenomenology and the Cognitive Sciences*, 13: 25-46.

Pacherie, E. (2017). "Collective phenomenology." In K. Ludwig & M. Jankovic (eds.), *The Routledge Handbook on Collective Intentionality*. New-York: Routledge, 162–173.

Pauer-Studer, H. (2014). "A Constitutive Account of Group Agency." *Erkenntnis* 79: 1623– 39.

Roth, A. S. (2003). "Practical Intersubjectivity." In F. Schmitt (ed.), *Socializing Metaphysics: the Nature of Social Reality*. Lanham, MD: Rowman & Littlefield, 65–91.

Rovane, C. (2014). "Group agency and individualism." *Erkenntnis*, 79 (9): 1663–1684.

Schroeder, M. (2007) *Slaves of the Passions*. Oxford: Oxford University Press.

Schroeder, M. (2012) "Precis of *Slaves of the Passions.*" *Philosophical Studies*, 152:431-434.

Schroder, T., Arpaly, N. (1999) "Alienation and Externality." *Canadian Journal of Philosophy*. 29 (3): 371-387.

Searle, J. (1990). "Collective Intentions and Actions." In P. Cohen, J. Morgan, and M. Pollack (eds.), *Intentions in Communication*. Cambridge: MIT Press, 401–415.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*. 10: 70–76.

Seemann, A. (2009). "Joint agency: Intersubjectivity, sense of control, and the feeling of trust." *Inquiry: An Interdisciplinary Journal of Philosophy*. 52: 500–515.

Szanto, T., Moran, D. (2015). "Introduction: Empathy and Collective Intentionality—The Social Philosophy of Edith Stein." *Human Studies,* 38: 445–461.

Tollefsen, D. (2002a) "Collective Intentionality and the Social Sciences." *Philosophy of the Social Sciences*, 32 (1): 25-50.

Tollefsen, D. (2002b). "Organizations as true believers". *Journal of Social Philosophy* 33 (3): 395–410.

Tollefsen, D. (2015). *Groups as agents*. Cambridge: Polity.

Trevarthen, C. (1980). The foundations of intersubjectivity: Development of interpersonal and cooperative understanding in infants. In D. R. Olson (Ed.), The social foundations of language and thought: Essays in honor of Jerome S. Bruner (pp. 316–342). New York: W. W. Norton.

Trevarthen, C. (2011). "The generation of human meaning: How shared experience grows in infancy." In A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (p. 73–113). MIT Press.

Tuomela, R. (2006). "Joint Intention, We-Mode and I-Mode". *Midwest Studies in Philosophy* 30 (1): 35-58.

Wilson, J. (2015). "Metaphysical Emergence: Weak and Strong." In T. Bigaj, C. Wuthrich (eds) *Metaphysics in Contemporary Physics. Poznan Studies in the Philosophy of the Sciences and Humanities*. 251-306.

Wilson, J. (2011). "Non-reductive Realization and the Powers-based Subset Strategy." *The Monist*. 94 (1): 121-154.

Wilson, J. (1999). "How Superduper does a Physicalist Supervenience Need to Be?" *The Philosophical Quarterly*. 49: 33-52.

Watson, G. (2004) *Agency and Answerability: Selected Essays*. New York: Oxford University Press.

Zaadnoordijk, L., S. Hunnius, M. Meyer, J. Kwisthout and I. van Rooij, (2015) "The developing sense of agency: Implications from cognitive phenomenology," *Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, Providence, RI, 2015, pp. 114-115, doi: 10.1109/DEVLRN.2015.7346126.