

# Shared Agency without Shared Intention

Samuel Asarnow  
Macalester College  
sasarnow@macalester.edu

Forthcoming in *The Philosophical Quarterly*

## Abstract

The leading reductive approaches to shared agency model that phenomenon in terms of complexes of individual intentions, understood as plan-laden commitments. Yet not all agents have such intentions, and non-planning agents such as small children and some non-human animals are clearly capable of sophisticated social interactions. But just how robust are their social capacities? Are non-planning agents capable of shared agency? Existing theories of shared agency have little to say about these important questions. I address this lacuna by developing a reductive account of the social capacities of non-planning agents, which I argue supports the conclusion that they can enjoy shared agency. The resulting discussion offers a fine-grained account of the psychological capacities that can underlie shared agency, and produces a recipe for generating novel hypotheses concerning why some agents (including, arguably, the great apes) do not engage in shared agency.

## Introduction

Humans and other agents interact with each other in many ways: strategically, adversarially, and cooperatively. Of particular interest to social scientists and philosophers are types of interactions dubbed ‘shared agency’ or ‘joint agency,’ as when agents walk together, dance a tango together, or build a house together.<sup>1</sup>

The constructivist tradition seeks to understand shared agency via psychological modeling.<sup>2</sup> Constructivists catalogue the features of different social interactions, and describe sets of mental states such that, if agents were in those mental states, they would be poised to engage in

---

<sup>1</sup> For these examples, see, e.g., (Bratman, 1992; Gilbert, 1990; Kutz, 2000).

<sup>2</sup> The label “constructivist” is inspired by Bratman; see, especially, (Bratman, 2014, pp. 30–35). Other approaches to shared agency include the influential “normativist” approach pioneered by Margaret Gilbert. See, e.g., (Gilbert, 2009). I set those approaches aside here, as my focus is on constructivism. On constructivism as modeling, see (Godfrey-Smith, 2005).

social interactions with those features. Constructivism promises a reductive view of shared agency, according to which its features can be understood without positing irreducible group agents.<sup>3</sup>

Constructivism's roots are in game theory, which models social interactions in terms of psychological posits representable by payoff functions, probability functions, and strategies. Game-theoretic solution concepts such as the Nash equilibrium are informative models of important types of agent sociality, such as the strategic interaction of agents in a prisoner's dilemma. Yet it is widely thought that game-theoretic constructions cannot model the distinctive features of shared agency.<sup>4</sup>

Constructivists thus often enrich the game-theoretic psychological palette, most notably by building psychological models that include intentions as well as beliefs and other mental states. Intention constructivism is arguably the leading approach to shared agency today.<sup>5</sup> Michael Bratman, an influential intention constructivist, has conjectured that the plausibility of intention constructivism supports what he calls the 'primacy of intention for modest sociality': intention is the key mental state to focus on when studying small-scale cases of shared agency.<sup>6</sup>

While I'm sympathetic to intention constructivism, I want to highlight here an underappreciated feature of the view: intention constructivists typically do not understand intentions simply as goal states or the outputs of decisions. Rather, following Bratman, they conceive of intentions as elements in (partial) plans, persisting states of mind whose elaborate functional roles underlie the capacity for sophisticated forms of temporally-extended agency, especially temporally-extended agency that involves complex future planning. Intention

---

<sup>3</sup> Compare (Alonso, 2018).

<sup>4</sup> See, e.g., (Bratman, 2014, pp. 5–6; Tuomela, 2009). For attempts to model shared agency in broadly game-theoretic terms, see, e.g., (Chant & Ernst, 2007; Ernst & Chant, 2007)

<sup>5</sup> A sample of influential versions of intention constructivism would be: (Alonso, 2016; Bratman, 2014; Kutz, 2000; McMahan, 2005; Miller, 2004; Searle, 1990; Tuomela, 2005; Velleman, 2000).

<sup>6</sup> As at, e.g., (Bratman, 2014, p. 29).

constructivists thus typically tie together the capacity for complex forms of sociality (such as shared agency) to the capacity for complex forms of temporally extended agency, or ‘planning agency.’

This connection between shared agency and temporally extended agency is striking, because (as some intention constructivists allow) not all agents are planning agents.<sup>7</sup> While it is plausible that many adult humans are planning agents, some children, adults with certain neurological conditions, and non-human animals are not. Yet they may enjoy complex social lives.

How sophisticated can the social lives of non-planning agents be? Can they engage in shared agency? Here the constructivist tradition offers little insight. Intention constructivism is silent, as these agents lack intentions. Game-theoretic constructivism has much to say about non-planning agents, but arguably it cannot distinguish shared agency from mere strategic interaction. And, as I will discuss below, the ‘minimalist’ forms of constructivism that have been developed to study the sociality of infants are too minimal to model the capacities of many non-planning agents.<sup>8</sup>

The goal of this paper is to fill this lacuna by developing a version of constructivism that is able to more adequately study the social lives of non-planning agents. My method is a Gricean ‘creature construction.’<sup>9</sup> I construct a series of hypothetical agents, beginning with simple belief-desire agents and gradually introducing further non-intention psychological resources. At each stage I catalogue the forms of sociality the agents might enjoy.

I argue for two conclusions. The first is a parity thesis: if agents who realize the kinds of constructions that Bratman and others emphasize count as engaging in shared agency, then so do

---

<sup>7</sup> Bratman calls this the “multiplicity of agency” thesis at (Bratman, 2018c, p. 15).

<sup>8</sup> I discuss minimalism in Section 3.

<sup>9</sup> See originally (Grice, 1974), as well as (Bratman, 2007).

certain non-planning agents. The kind of sociality non-planning agents can in principle enjoy differs from the sociality of planning agents only in not being temporally extended. It is no less ‘social’ or ‘shared.’

The second conclusion is a corollary of the parity thesis: shared agency does not require temporally extended agency. While it is plausible that many adult humans make use of the same psychological capacities for temporally extending their agency as for socially extending it (as it were), the capacities are in principle separable. This highlights the importance for social ontologists of thinking in a fine-grained way about the various capacities that underlie shared agency, other than intention. As I articulate below, these include what I call the ‘taste for sociality,’ ‘cognitive-conative’ theory of mind, and the acceptance of certain norms of rationality.

The paper proceeds as follows. Section 1 describes the constructivist approach to social ontology and explains a key assumption of my argument, pluralism about social ontology. Section 2 introduces Bratman’s intention constructivism. Section 3 discusses the characteristics of the non-planning agents I intend to study. Sections 4-6 construct a quartet of non-planning agents, argue for the parity thesis, and reply to an objection. Section 7 draws from this discussion three lessons for social ontologists, and considers how my argument might inform debates about the social capacities of great apes and other non-human agents.

## **1. Shared Agency, Constructivism, and Pluralism about Social Ontology**

Constructivists seek to understand agent sociality by constructing psychological models. Of special interest are forms of sociality dubbed shared agency.

Shared agency is often introduced by contrast with merely parallel agency.<sup>10</sup> To take a much used pair of ‘contrast cases,’ consider two cases involving two people walking together down Fifth Avenue in New York City.<sup>11</sup> In one case, the agents are taking a walk together. In the second case, the agents are strangers, aware of each other’s presence and maintaining appropriate distance from each other in light of that awareness, but not interacting in any further way. For a period of time, the two pairs might be indistinguishable from the outside: the members of each pair might be walking at the same orientation to each other, at the same distance from one another, and so on. Yet there is an intuitive difference between the two pairs: there is a straightforward sense in which the first pair are ‘walking together’ while the second are not. ‘Shared agency’ names what one pair exhibits and the other does not.

Constructivists begin studying shared agency by cataloguing its features, before constructing psychological models realizing those features. Highly influential is Bratman’s feature catalogue. According to Bratman, paradigm cases of shared agency involve mental states that play three functional roles.<sup>12</sup> First, the agents’ mental states coordinate each individual’s actions as they engage in the shared activity. That is, they guide each agent’s behavior, so that each tracks the shared goal of the agents acting together, monitoring the other’s behavior and adjusting as appropriate. Second, the agents’ mental states provide a background against which the agents can resolve disagreements, bargain, and make decisions concerning new options that arise in the course of the shared action. Finally, the mental states motivate the agents to plan together about how their activity will evolve over time, and about the division of labor and roles, if need be.

---

<sup>10</sup> Compare, e.g., (Blomberg, 2015a, p. 169).

<sup>11</sup> See (Bratman, 2014, p. 5), following (Gilbert, 1990), among others.

<sup>12</sup> See, e.g., (Bratman, 2014, pp. 27–29). My presentation follows (Alonso, 2016, p. 276).

Drawing on this kind of characterization, Bratman and others have argued that while parallel action can be modeled game-theoretically, shared agency cannot.<sup>13</sup> But they argue that positing additional types of mental states, notably intentions, allows constructivists to model shared agency.

Before giving an example of such an intention constructivism, a methodological comment is in order. I doubt that there is a unique phenomenon with clearly defined boundaries that deserves the name shared agency. Instead, I suspect that there are many different kinds of sophisticated social interactions worth of study, including social interactions that lack some of the three features just mentioned, as well as interactions that involve further features. Different agents, and different kinds of agents, may engage in different forms of sophisticated social interactions across their lifespans. Many of them may be worthy of study and can correctly be referred to as shared agency.

Loosely following Cédric Paternotte and Stephen Butterfill, I call this idea pluralism about the target of theorizing in social ontology, or pluralism about social ontology for short.<sup>14</sup> Commitment to pluralism has not often been made explicit in the literature. There are signs that some philosophers would accept it, however. Bratman, for example, has studied several different types of social interactions, including shared agency and shared cooperative activity, indicating that he might be sympathetic to pluralism about the target of theorizing in this area.<sup>15</sup> And, as I will discuss below, he also occasionally gives examples of cases of shared agency that lack one or more of the three characteristics described above.<sup>16</sup> Other social ontologists who have indicated that they accept something like this pluralism include Olle Blomberg, Stephen Butterfill, and

---

<sup>13</sup> Again, see (Bratman, 2014, pp. 5–6; Tuomela, 2009).

<sup>14</sup> See (Butterfill, 2016, p. 14; Paternotte, 2015).

<sup>15</sup> Bratman, identifies as objects of study strategic interaction, shared agency, and shared cooperative activity (in (Bratman, 1992)), and has discussed other constructions with different features in unpublished work.

<sup>16</sup> As when Bratman imagines cases of spontaneous and temporary shared agency; see (Bratman, 2006, p. 7)

Christopher Kutz.<sup>17</sup> (I foreground my commitment to pluralism here because the interest of my eventual argument will depend on it.)

Pluralists may suspect that participants in the literature on shared agency have sometimes be talking past one another (and characterizing different kinds of social interactions). But pluralism is compatible with seeing plenty of genuine disagreement. For example, it is important to debate which properties certain actual or hypothetical cases of social interaction tend to have, and how best to model social interactions with those properties.<sup>18</sup>

That being I will rely on Bratman's three part characterization as a guide below, though I will not view its features as necessary for shared agency.

## **2. Intention Constructivism and the Primacy of Intention Thesis**

There are many different intention constructivisms in the literature, but I will model my creature constructivism in sections 3 and 4 on Bratman's account. So in this section I will briefly explain how Bratman conceives of intention, and what his intention constructivism looks like.

### *2.1 Bratman on Intention*

For Bratman, intentions are motivational mental states that constitute plan-like commitments to act.<sup>19</sup> Intentions play three distinctive functional roles.

First, intentions settle practical questions. This means one who forms an intention to  $\phi$  ceases deliberating about whether to  $\phi$ , and believes or is disposed to take for granted that they

---

<sup>17</sup> See (Butterfill, 2015, p. 152, 2016, p. 14; Kutz, 2000, p. 4), the argument against "dualism" in (Blomberg, 2015a), and, in a way, (Tuomela, 2005, sec. 4).

<sup>18</sup> This may be part of what normativists and constructivists disagree about. Compare (Alonso, 2016).

<sup>19</sup> See (Bratman, 1987), which I follow in this section.

will  $\phi$ . It also means that intentions are stable: they persist in ways that permit planning agents to resist certain temptations.<sup>20</sup>

Second, intentions filter out inconsistent plans. This means that one who intends to  $\phi$  is disposed to ignore believed alternatives to  $\phi$ -ing, to rule them out if they do become salient, and (especially) not to intend believed alternatives.

Finally, intentions pose means-ends problems. One who intends to  $\phi$  will deliberate about how to  $\phi$  and will form intentions to take believed necessary means. This is in part what allows planning agents to engage in temporally-extended activities with a complex hierarchical structure: their intentions dispose them to discover and fill in the gaps in their plans, leading to the interlocking and mutually reinforcing sets of plans and subplans characteristic of temporally-extended agency.

Each functional role is associated with a norm or requirement of rationality. These requirements enjoin stability of intention, consistency or agglomeration of intentions, and means-ends coherence.<sup>21</sup> Moreover, planning agents accept these norms of rationality, which entails (inter alia) that they are stably (albeit defeasibly) disposed to conform with them. This norm acceptance is understood (roughly) in terms of the possession of a general intention to conform with these norms.<sup>22</sup>

---

<sup>20</sup> On this point, see the essays in (Bratman, 2018c).

<sup>21</sup> Here too see (Bratman, 2018c).

<sup>22</sup> See, e.g., (Bratman, 2014, pp. 16–17).



## 2.2 Bratman's Intention Constructivism

Now turn to Bratman's construction. Bratman offers sufficient conditions for shared agency. He claims that two agents who instantiate a complex set of psychological properties, each of whom is J-ing, count as J-ing together, as a shared action.<sup>23</sup>

### Planning Agent Sociality

- (1) Each agent intends: (a) that both agents J, (b) by way of each's intention that both J, and (c) by way of sufficiently meshing subplans.
- (2) The agents' intentions are (a) persistence interdependent and (b) mutually responsive
- (3) All of this is common knowledge.

Each clause requires clarification. About (1), three comments. First, to avoid circularity, the intention that we J in (1)(a) is understood in a 'distributed' sense: it is an intention that both J, in the sense in which two agents might J in parallel (i.e., not engaged in shared agency). Because this involves each agent intending that the other J, it involves a disposition to seek means not only to their own J-ing, but also to the other's J-ing. This is a disposition to help or assist that is characteristic of shared agency. Second, clause (1)(b) rules out 'mafia cases,' in which one agent intends that the other agent J, but plans to force them to J in a way that bypasses the second agent's agency (e.g., by kidnapping them). Finally, (1)(c) ensures that each agent is committed to filling out the as-yet-undecided details of their plans in compatible ways. This rules out cases in which the agents both intend that they J, but one agent plans to 'override' the other's subplans should they conflict with the agent's own subplans. This rules out certain cases of manipulation and deception, provides a framework for future bargaining and shared decision-making, and renders the shared action distinctively durable.

---

<sup>23</sup> Throughout this section, I simplify the discussion in (Bratman, 2014, Chapters 2–3)

About (2), two comments. First, the agents' intentions are persistence interdependent in the sense that each continues to intend that both J if and only if they believe the other continues so to intend.<sup>24</sup> This allows Bratman to avoid the objection that I cannot intend your actions; given that I believe (2)(a), obtains, I believe I can settle whether you J, and so my intention that you J satisfies David Velleman's 'settle condition.' Finally, (2)(b) ensures that the agents monitor and respond to each other while acting, ruling out 'prepackaging' cases where the agents in fact proceed in parallel despite earlier joint planning.

Finally, (3) ensures that all of this is 'out in the open.' Both agents are aware of what they are doing, and of the various conditions that each satisfies.<sup>25</sup>

Bratman argues convincingly that agents engaging in Planning Agent Sociality have mental states that play the three functional roles associated with shared agency. Their mental states dispose them to coordinate their behavior so as to track the shared goal, frame shared bargaining and decision-making, and support future planning concerning how the shared activity will proceed.

Importantly, Bratman argues that the plausibility of his intention constructivism supports a thesis he calls 'the primacy of intention for modest sociality.'<sup>26</sup> This thesis is that intention is the key attitude to use in our theorizing in social ontology. It is not entirely clear to me how Bratman intends this thesis to be understood. He often mentions this thesis in the context of discussing whether shared agency requires convergence in evaluative belief, which he denies. So, on a weak reading, the thesis is that intention is more primary than evaluative belief to our understanding of shared agency. By contrast, a strong version of this thesis would hold that intention is primary over

---

<sup>24</sup> While Bratman suggests that there can be shared agency without persistence interdependence of intention, he thinks persistence interdependence is an important feature of many cases of shared agency. See (Bratman, 2014, p. 77). I emphasize it here because (as I argue in Section 4) non-planning agents can enjoy something like it.

<sup>25</sup> Arguably, sufficient conditions for shared agency can be given without appeal to (3); see (Blomberg, 2016).

<sup>26</sup> See, e.g., (Bratman, 2014, p. 29).

other motivational states, as well—that intention, not (say) desire is the key to understanding shared agency. I don't know whether Bratman would accept the strong version of this thesis (and I suspect he would not), but it is interesting to consider whether it is correct. I turn to that question in the remainder of the paper.

### **3. The Multiplicity of Agency and the Psychology of Non-Planning Agents**

It may well be that when planning agents engage in shared agency, they realize something like Planning Agent Sociality. Yet not all agents are planning agents: some agents engage in purposive behavior, tracking and pursuing goals over relatively short periods of time, but lack intentions. These agents lack the capacity for sophisticated, hierarchical planning and thus temporally-extended agency, but they may have memories, expectations for the future, and even a temporally-extended self-concept. Plausibly, small children, many non-human animals, and adults with certain neurological disorders are non-planning agents. Importantly, this is acknowledged even by theorists of planning agency. Bratman, for example, refers to the existence of non-planning agents as an aspect of the 'multiplicity of agency.'<sup>27</sup> Arguably, even planning agents need not make use of their planning capacities at all times. One can imagine an agent who is sometimes motivated by Bratmanian intentions, but at other times (say, when acting spontaneously) is motivated and guided only by beliefs and desires.<sup>28</sup>

Non-planning agents immediately raise questions of social ontology: what kinds of sociality can they engage in, and can they engage in genuine shared agency? These questions have attracted almost no attention in the literature.

---

<sup>27</sup> As at, e.g., (Bratman, 2018b, pp. 5–6). Bratman variously mentions squirrels (Bratman, 2018a, pp. 58–59) and dogs and cats (Bratman, 2014, p. 23) as examples of non-planning agents.

<sup>28</sup> Though Bratman would reject this idea. See (Bratman, 1987, Chapter 8).

There is one constructivist research program that has addressed a similar issue. Minimalist constructivists have highlighted that infants and very young toddlers have important social capacities (such as the capacity for ‘joint attention’ with caretakers) despite lacking intentions, theory of mind, and other psychological capacities presupposed by intention constructivists.<sup>29</sup> They have thus constructed models of agent sociality that begin with ‘the simplest possible starting point.’<sup>30</sup>

Minimalist constructivism is of great interest, and I will briefly compare my approach to a minimalist approach further below, but it is not the final word about the social lives of non-planning agents. There are many interesting (actual and hypothetical) non-planning agents whose mental capacities lie between those of infants and adult humans, including agents with various forms of theory of mind but without intentions.<sup>31</sup> Minimalism does not address those agents. There is thus a gap in the literature concerning whether agents more sophisticated than infants but less sophisticated than adult humans can engage in shared agency. This is a striking gap, because I suspect that many actually existing creatures are non-planning agents of roughly this sort.

Before attempting to fill that gap, I want to comment on how I understand the minds of the relevant agents. I will model these agents in terms of beliefs and desires, as distinct from intentions.<sup>32</sup> Five points about desires are important. First, desires do not settle deliberative questions. Desiring an outcome does not dispose one to cease deliberating about whether to pursue that outcome, and desires are not characteristically stable over long periods of time (although they may happen to be stable in any particular case). Second, desires do not filter out inconsistent

---

<sup>29</sup> For minimalist approaches, see, e.g., (Blomberg, 2011, 2015b, 2015a; Butterfill, 2012, 2016; Butterfill & Sebanz, 2011; Elisabeth Pacherie, 2011, 2013; Tollefsen, 2005; Vesper, Butterfill, Knoblich, & Sebanz, 2010).

<sup>30</sup> See (Butterfill, 2016, p. 4).

<sup>31</sup> For discussions of the minimalist emphasis on agents without sophisticated theory of mind see (Butterfill, 2012, pp. 41–42; Tollefsen, 2005, p. 81).

<sup>32</sup> In emphasizing desires rather, my approach resembles that of (Blomberg, 2015a), though his focus is on agents without theory of mind.

options. It is possible (and common) to desire two believed incompatible outcomes, and desiring one outcome does not dispose one to ignore incompatible outcomes. Third, desires do not pose means-ends problems. An agent may desire an outcome without deliberating about necessary means, and without desiring believed necessary means (though, in a particular case, an agent may engage in such deliberation and may form a desire for believed necessary means). Fourth, desires are not subject to requirements of rationality.<sup>33</sup> Finally, I will assume that desires come in various degrees of strength, that (ceteris paribus) non-planning agents to act on their strongest desires,<sup>34</sup> and that desire-belief motivation can be sufficient for intentional action.<sup>35</sup>

In Section 5 I will introduce creatures who have volitions as well as beliefs and desires. A volition, in my terminology, is a present-directed or proximal intention: an intention to act now or in the very near future.<sup>36</sup> Forming a volition does close a deliberative question (since it is the result of deciding), and volitions are stable over short periods of time, but they do not make possible the kind of temporally-extended agency that Bratmanian intentions do. Volitions, unlike desires, are subject to a requirement of rationality that links them to desires and beliefs. I will discuss volitions in more depth in Section 5.<sup>37</sup>

---

<sup>33</sup> For two of this assumption defenses, see (D. Baker, 2010; Bratman, 1987, Chapter 1). If desires are subject to requirements of rationality, however, that would only strengthen the case for the parity thesis; compare Section 5.2.

<sup>34</sup> Note that this assumption is weaker than the assumption I make in (Asarnow 2019).

<sup>35</sup> As in, e.g., (Davidson, 1980). As I discuss in Section 6, when an agent acts on what they most desire, that desire plays the same role as an “intention-in-action” in the sense of Tollefsen (following Searle). Compare (Blomberg, 2011; Searle, 1983; Tollefsen, 2005, pp. 88–89).

<sup>36</sup> Compare (Mele, 1992, Chapters 8, 10).

<sup>37</sup> One minimalist who has appealed to something like volitions is Elizabeth Pacherie, in (Elizabeth Pacherie, 2013). See also her discussion of “intermediate levels of planning” in (E Pacherie, 2007).

#### **4. Creature Construction, Part 1: the Taste for Sociality and Cognitive-Conative Theory of Mind**

My method in studying the social lives of non-planning agents will be a Gricean creature construction.<sup>38</sup> Here and in Section 6, I describe a series of agents of increasing cognitive sophistication, and study the social possibilities that each agent's mind affords. The goal of the process will be to see how closely non-planning agents can approximate the kind of interactions characteristic of Planning Agent Sociality, with close attention paid to what social possibilities are offered at different waypoints along the journey. Like Bratman, my intention is to provide sufficient conditions only.

Begin with a creature who has beliefs and desires, but whose mental states take only a restricted range of contents.<sup>39</sup>

Creature 1     This agent has beliefs and desires, but lacks theory of mind. It has beliefs about the behavior of other creatures, but it does not have desires about the behavior of other creatures.

This is the kind of agent studied in elementary game theory, where the payoff of an outcome for an agent is taken to be a function of years in prison, or dollars, rather than a function of the other agent's behavior. Creature 1 can be studied fruitfully using game theoretic constructions, including that of the Nash equilibrium.<sup>40</sup>

Taking Nash equilibrium as our cue, we can model an important case of the sociality of Creature 1 agents as follows.<sup>41</sup> When the two creatures are J-ing:

---

<sup>38</sup> See, e.g., (J. Baker, 2008; Bratman, 2007, 2014; Grice, 1974; Railton, 2014). Compare also (Godfrey-Smith, 2005).

<sup>39</sup> In particular, both its ultimate (or intrinsic) and its instrumental desires concern only states of affairs not specified in terms of the behavior of others. On that distinction, see, e.g., (Sober & Wilson, 1998, pp. 217–222).

<sup>40</sup> Though if strategies are psychologically realized by intentions, many solution concepts cannot be realized by Creature 1 agents.

<sup>41</sup> Compare (Bratman, 2014, p. 41).

## Creature 1 Sociality

- (1) Each creature truly believes that both creatures are J-ing.
- (2) Each creature most desires that it J, given that the other creature is J-ing.

That is, each creature believes (what is true) that each creature is doing its part in their J-ing, understood distributively. And each creature's desire to continue doing its part in their J-ing is stronger than any desires for outcomes it believes incompatible with that.

Creature 1 Sociality is a far cry from shared agency. Many cases of merely parallel action may be of this type, including parallel defection in a prisoner's dilemma.

One way to enrich the social capacities of these agents is to allow their desires to take a broader range of contents.<sup>42</sup> We thus turn to Creature 2.

**Creature 2** This agent has beliefs and desires, but lacks theory of mind. It has both beliefs and desires concerning the behavior of other creatures.

This agent has desires about how other agents act, and is thus in a position to desire that it J with another creature (in the distributed sense). If such a creature in fact has such desires, I say it has the 'taste for sociality.' Creature 2 agents with this taste can realize the following:

## Creature 2 Sociality (Type 1)

- (1) Each creature truly believes that both creatures are J-ing.
- (2) Each creature most desires that both creatures J.

These agents not only most desire to do what they are doing, but they also desire that the other creature be doing what it is doing.<sup>43</sup>

---

<sup>42</sup> There may be other ways to proceed here, for example, by turning clause (2) into a biconditional.

<sup>43</sup> The desires specified in (2) may be either ultimate or instrumental. Perhaps an agent whose relevant desires are ultimate is capable of more deeply social agency than one whose desires are instrumental. But paradigmatic intention

The taste for sociality is highly consequential: it makes possible two characteristic features of Planning Agent Sociality. First, these agents may (though are not guaranteed to) engage in the kind of helping behavior described in section 2.2 above. The desire for an outcome may cause a desire for a believed necessary means in a particular case, and so one agent's desire that the other J may lead to desires for believed necessary means to the other agent J-ing—for example coaxing, urging, or removing obstacles.

Second, it makes possible a first approximation of the idea of persistence interdependence, as follows.

#### Creature 2 Sociality (Type 2)

- (1) Each creature truly believes that both creatures are J-ing.
- (2) Each creature most desires that (a) both creatures J, (b) and that each Js if and only if the other Js.

Such creatures' behavior approximates persistence interdependence, differing from Bratmanian persistence interdependence in two ways. First, Bratmanian persistence interdependence arises by way of each agent's acceptance of requirements of rationality. By contrast, here it arises by way of the causal efficacy of the desire specified in (2)(b).<sup>44</sup> Second, Bratman suggests that the persistence of each agent's intentions depend not only on the continued behavior of the other agent, but also on the continued intending of the other agent. Since Creature 2 lacks theory of mind, it cannot represent the other creatures' mental states in an appropriate way.<sup>45</sup>

---

constructivists allow for (as it were) purely instrumental shared agency. So establishing my parity thesis does not require appeal to ultimate desires here. Compare, e.g., (Bratman, 2014, pp. 37–39).

<sup>44</sup> Compare (Bratman, 2014, p. 65). I emphasize desired persistence interdependence while Bratman emphasizes believed and actual persistence interdependence. I address this difference when I introduce the idea of a “success case” in Section 5.1.

<sup>45</sup> Notably, Creature 2 Sociality (Type 2) importantly resembles certain minimalist constructions. Compare Butterfill's concept of a “shared goal,” which “coordinate [the agents'] goal-directed activities around an outcome to be achieved as a common effect of their efforts” (Butterfill, 2012, p. 37).



The third creature we will consider thus has all of the capacities of Creature 2, as well as what I call ‘cognitive-conative’ theory of mind. Many discussions of theory of mind emphasize the capacity to form beliefs about the mental states of others.<sup>46</sup> Yet many agents not only have beliefs about the mental states of others, but also desires—they desire that other agents’ minds be certain ways. I call the capacity to have beliefs and desires about other agents’ minds cognitive-conative theory of mind, and the capacity to have beliefs (but not desires) about other agents’ minds (merely) cognitive theory of mind. Hence Creature 3:

Creature 3     This agent has beliefs and desires and has cognitive-conative theory of mind.

Such a creature can thus have both beliefs and desires about the beliefs and desires of other agents.

Creature 3 allows for the introduction of five innovations that bring its sociality more closely in line with Planning Agent Sociality. Two are closely related and can be introduced at once.

Creature 3 Sociality (Type 1)

- (1) Each creature truly believes that both creatures are J-ing.
- (2) Each creature most desires that (a) both creatures J (b) by way of their desires to J, and (c) that each Js if and only if the other Js by way of its desire to J.

Clause (2)(c) allows for a kind of persistence interdependence that more closely approximates Bratman’s. And the new (2)(b) rules out ‘mafia cases,’ in which the participation of one creature does not go by way of its mental states and thus is not intentional.<sup>47</sup>

---

<sup>46</sup> For a summary of the development of research on cognitive theory of mind, see (Andrews, 2016, sec. 4.2).

<sup>47</sup> Arguably, the capacity to distinguish intentional from non-intentional behavior does not require cognitive theory of mind. If so, Creature 2 agents can approximate (2)(b). For discussion of this issue, see (Tomasello, Carpenter, Call, Behne, & Moll, 2005), as well as (Butterfill, 2012). Butterfill, Pacherie, and Tollefsen make this assumption in (Butterfill, 2012; Elisabeth Pacherie, 2013; Tollefsen, 2005).

Next, we can introduce a condition that parallels Bratman's idea of mutual responsiveness.

### Creature 3 Sociality (Type 2)

- (1) Each creature truly believes that both creatures are J-ing.
- (2) Each creature most desires that (a) both creatures J (b) by way of their desires to J, and (c) that each Js if and only if the other Js by way of its desire to J.
- (3) Each creature most desires that their continued J-ing be responsive to the desires and actions of the other.<sup>48</sup>

This condition rules out prepackaged shared action.

Fourth, theory of mind allows us to approximate meshing subplans. While non-planning agents cannot intend to mesh their subplans, they may well have desires that would play an analogous role, viz., to rule out certain forms of manipulation and deception, provide a framework for shared decision-making, and enhance stability. For example:

### Creature 3 Sociality (Type 3)

- (1) Each creature truly believes that both creatures are J-ing.
- (2) Each creature most desires that (a) both creatures J (b) by way of their desires to J, and (c) that each Js if and only if the other Js by way of its desire to J.
- (3) Each creature most desires that their continued J-ing be responsive to the desires and actions of the other.
- (4) Each creature most desires that each agent does not do anything incompatible with what the other agent most desires.<sup>49</sup>

Here, clause (4) rules out (for example) cases in which one agent desires to bypass the other agent's chosen way of J-ing, and promotes problem-solving via bargaining rather than coercion or deception.

---

<sup>48</sup> This almost a direct quotation from (Bratman, 1992, p. 94).

<sup>49</sup> On how to understand the relevant sense of "incompatible," see, e.g., (Núñez, 2019).

Finally, theory of mind allows for all of these conditions to be, as Bratman puts it, ‘out in the open.’ I call this Creature 3 Sociality, simpliciter:

### Creature 3 Sociality

- (1) Each creature truly believes that both creatures are J-ing.
- (2) Each creature most desires that (a) both creatures J (b) in part by way of their desires to J, and (c) that each Js if and only if the other Js in part by way of its desire to J.
- (3) Each creature most desires that their continued J-ing be responsive to the desires and actions of the other.
- (4) Each creature most desires that each agent does not do anything incompatible with what the other agent most desires.
- (5) That (1)-(4) obtain is common true belief.

Here ‘common true belief’ is an analogue of common knowledge that can be instantiated by true beliefs that may not constitute knowledge. Each agent believes that the other agent believes that (1)-(4) obtain, and they believe that the other agent believes that, and so on, supporting their joint decision-making and bargaining.<sup>50</sup>

Together, the taste for sociality and cognitive-conative theory of mind allow for analogues of each feature of Planning Agent Sociality. Clauses (1)(a) and (1)(b) of Planning Agent Sociality are approximated by clause (2) of Creature 3 Sociality; clause (1)(c) is approximated by clause (4); clause (2)(a) is approximated by (1)(c), and (2)(b) by (3); and clause (3) by (5).

I will shortly argue that instantiating (1)-(5) is sufficient for a kind of social interaction that constitutes shared agency. I do not claim that these conditions are necessary. Shared agency among

---

<sup>50</sup> Two comments. First, for Bratman, a central role of the common knowledge condition is to support hierarchical planning. If the kind of bargaining Creature 3 agents can participate in does not rely on common true belief, then condition (5) may be idle. Compare (Bratman, 2014, pp. 57–59). For an argument that the common knowledge condition is idle even in Bratman’s construction, see (Blomberg, 2016). Second, arguably, obtaining evidence for these beliefs (especially that the other agent most desires something) would require additional capacities above those specified here—e.g., certain communication capacities. I am not sure whether this is so, but it is plausible that such capacities do not require full-blown planning agency, and so non-planning agents could in principle enjoy them.

Creature 3 agents may be multiply realizable, or may be realizable by agents instantiating only some of these conditions.<sup>51</sup>

## **5. The Parity Thesis and the Precariousness Objection**

Imagine two Creature 3 agents walking down Fifth Avenue, instantiating Creature 3 Sociality. Each most desires to walk the way it is walking, given that the other walks the way it is walking, and each most desires that the other walk in the way it is walking, by way of its desire to do so. Each desires that their activity and desires be persistent interdependent, in that each desires that each continue walking if and only if the other continues walking by way of its desire to walk. Each desires that their motivations and behavior be mutually responsive. Each desires that each not do anything incompatible with what the other most desires. All of this is out in the open.

These agents closely approximate agents engaged in Planning Agent Sociality. But how close an approximation is it? I consider that question in this section, and in so doing provide my central argument for the parity thesis (to which I make two small revisions). I then turn to an objection to the parity thesis, which I will address in Section 6.

### *5.1 The Argument for the Parity Thesis*

Two differences between Creature 3 Sociality and Planning Agent Sociality stand out. One difference, I argue, can be minimized by introducing a further idea. The other difference does not disqualify Creature 3 Sociality from counting as shared agency. So, I conclude, if Planning Agent Sociality is sufficient for shared agency, so is Creature 3 Sociality.

---

<sup>51</sup> Per Bratman, condition (3) may be unnecessary; per Blomberg, condition (5) may be unnecessary.

The first difference stems from a difference between desires and intentions. Because desires are not commitments to act, agents may realize Creature 3 Sociality without actually doing some of the things they desire to do. They may act on their desires not to do anything incompatible with what the other agent desires, but they may not. They may form desires to take the believed necessary means to their desires (and so, e.g., help each other), but they may not. Their desires may be stable over time, but they may not be. And so on. So agents might realize the mental states specified by Creature 3 Sociality without acting on the specified desires, and so without doing the things characteristic of shared agency.

This difference can be minimized. While non-planning agents do not always act on their desires, and do not always desire to take the believed necessary means to their desires, sometimes they do. In what I will call a ‘success case’ of Creature 3 Sociality, the agents in fact stably maintain their desire that they walk together over a period of time, form desires for believed necessary means, and act on all of the relevant desires. Importantly, this idea is not ad hoc. Bratman implicitly makes use of the idea of a success case when his construction appeals to actual persistence interdependence and mutual responsiveness (rather than merely believed or intended persistence interdependence and mutual responsiveness).

A revised version of the parity thesis would thus state that if agents realizing Planning Agent Sociality count as engaged in shared agency, so do agents realizing success cases of Creature 3 Sociality.

The second difference between this Creature 3 Sociality and Planning Agent Sociality is revealed when we consider whether agents in a success case of Creature 3 Sociality have mental states that playing three functional roles in Bratman’s initial characterization of shared agency.

Such agents clearly satisfy two of the three roles. First, their desires and beliefs coordinate their activity as it unfolds, specifying a shared goal at which they both aim (their J-ing) and disposing them to pursue that goal, monitor the other agent's progress toward that goal, and help the other agent. Second, their desires and beliefs provide a background against which the agents can bargain and make decisions: their desires to continue J-ing by way of their desires, and to continue acting in a way consistent with the other agent's desires.

However, the desires of Creature 3 agents do not play the third target role: they do not support future planning. Indeed, shared planning is in principle ruled out for these non-planning agents. This an important limitation: Creature 3 agents are in principle unable to plan a party together, let alone manage a corporation together or govern a state together.

Does this difference disqualify Creature 3 Sociality from counting as shared agency? I think it does not. Many paradigm examples of shared agency do not seem to involve intentions that play this functional role. This is particularly clear in cases of spontaneous and (intentionally) temporary shared agency. One example of such a case comes from Bratman:

Suppose you and I are independently walking down Fifth Avenue. We spot each other on 65th Street, and we briefly walk together, chatting, until, as it happens, you peel off at 59th Street. We do not merely walk individually along the same stretch of street and at the same time. Rather, we intentionally walk together for this brief time, and we briefly have a shared intention to walk together.<sup>52</sup>

Plausibly, these agents may be engaged in shared agency even if each walks with the other only on the condition that their walking together is temporary, that is, even if neither is set to plan together with the other.

---

<sup>52</sup> (Bratman, 2006, p. 7).

This example suggests that having states playing this third functional role is unnecessary for shared agency. From a pluralist standpoint, cases that involve shared planning and cases that do not are simply different forms of highly sophisticated sociality, one temporally-extended, and one not.

This completes my preliminary defense of the parity thesis. In its revised form, the parity thesis states that if cases of spontaneous, intentionally temporary Planning Agent Sociality count as shared agency, so do success cases of Creature 3 Sociality. The sociality of Creature 3 agents differs from the sociality of planning agents only in lacking temporal extension; it is (as it were) no less social.

## *5.2 The Precariousness Objection*

At this point, I want to consider an important objection to the parity thesis. This objection stems from the way that the idea of a success case was able to minimize, but not eliminate, the first difference between Creature 3 Sociality and Planning Agent Sociality. It claims that the behavior and motivations of Creature 3 agents are precarious in two ways that the behavior and motivations of planning agents are not, and that this grounds a significant difference between Creature 3 Sociality and Planning Agent Sociality

The first aspect of this precariousness has to do with the fickle nature of Creature 3 agents' desires in a success case. The agents' desires happen to have brought them to act, happen to have caused them to desire the believed necessary means to their desired ends, and happen to have been stable over a period of time. But none of that was the result of the playing out of the characteristic functional roles of those states. By contrast, when planning agents engage in Planning Agent

Sociality, their intentions durably motivate them to act, lead them to take means, and are stable over time as a result of their essential features.

The second aspect has to do with rationality. The intentions of agents engaged in Planning Agent Sociality implicate them in a web of rational requirements supporting the shared action. A planning agent who intends to do something they believe incompatible with J-ing is criticizably irrational, as is a planning agent who fails to intend the believed necessary means to the other agent's J-ing, and, arguably, a planning agent who fails to act on their relevant intentions.<sup>53</sup> By contrast, a Creature 3 agent who forms and acts on a desire to do something incompatible with J-ing, who fails to form the desire for a believed necessary means to the other agent's J-ing, or who fails to act on a relevant desire, is not criticizably irrational.

Planning agents engaged in Planning Agent Sociality are thus durably disposed and (as it were) rationally committed to elements of their joint actions, in a way that Creature 3 agents engaged in Creature 3 Sociality are not. One might hold that this difference shows that Creature 3 Sociality is not genuinely shared agency, since there is neither a motivational nor rational commitment to the shared action.

I am tempted to respond to this objection simply by pointing out that planning agents are sometimes irrational. Planning agents sometimes fail to act on their intentions, to intend necessary means, to filter out inconsistent plans, and so on. So there is an element of precariousness even in the motivations of planning agents. Yet it is not plausible that planning agents who are often irrational are thereby unable to (say) walk together. Because of this, I doubt there is much of a difference between some cases of Planning Agent Sociality and some success cases of Creature 3 Sociality.

---

<sup>53</sup> Though I raise a puzzle about this issue in (Asarnow 2019).



Another response is possible, however. Neither of the key features that this objection points to—durable dispositions to act, and webs of rational commitments—requires intentions in Bratman’s sense. There can be non-planning agents with mental states that feature both of those characteristics. I turn to such agents now.

## **6. Creature Construction, Part 2: Rationality without Temporal Extension**

A wide range of philosophers have discussed types of mental states that involve durable dispositions to act and constitute a kind of rational commitment, but that do not have the temporally-extended structure of intentions, as Bratman conceives them. These mental states include guiding desires, settled objectives, and proximal or present-directed intentions.<sup>54</sup> In what follows, I will consider agents with proximal intentions, which I will call ‘volitions’ to avoid confusion. However, the response to the precariousness objection could be developed by considering agents with any of these mental state types.

I will introduce volitions by contrasting them with both desires and intentions (in Bratman’s sense). Volitions are unlike desires in four ways. First, volitions are characteristically brought into existence by decisions—they are the outputs of decisions. Second, they are linked to behavior in a more durable, law-like way than desires (even strongest desires). To a first approximation, if an agent has a volition to  $\phi$  now and nothing prevents the agent from  $\phi$ -ing now,

---

<sup>54</sup> For “guiding desires,” see (Bratman, 1987, pp. 137–138). For “settled objectives,” see (McCann, 1991); Bratman suggests that squirrels may have settled objectives at (Bratman, 2018a, pp. 57–58). For proximal intentions, see, e.g. (Mele, 1992, Chapters 8, 10). Volitions in my sense should not be confused with the volitions discussed by volitionalists. Those are mental acts of trying, which are said to be a necessary means to any bodily action. I discuss volitions in that sense in (Asarnow 2019). Pacherie appeals to something like volitions at (Elisabeth Pacherie, 2013, pp. 1825–1836). Compare also (Butterfill, 2012, p. 39).

the agent will  $\phi$  now.<sup>55</sup> Third, volitions, unlike desires, are characteristically stable over short periods (as I discuss shortly).<sup>56</sup>

Finally, volitions are subject to a requirement of rationality relating them to beliefs and desires. How to understand this requirement is controversial. For simplicity, I will assume it is decision-theoretic in structure: when an agent with volitions deliberates, rationality requires that they decide upon the option that maximizes the expected satisfaction of their current desires.<sup>57</sup> Agents with volitions accept this norm of rationality, which entails (inter alia) that they are durably disposed to form volitions in accordance with it. And, arguably, there is an additional norm enjoining agents to act on their volitions.<sup>58</sup>

Volitions (as I will conceive of them) differ from intentions in four ways.<sup>59</sup> First, volitions are less stable than intentions, and there is no norm of rationality directly enjoining volition stability. However, the norm of expected desire-satisfaction maximization de facto enjoins stability, in that an agent with a rationally-formed volition to  $\phi$  who reopens deliberation about whether to  $\phi$ , but whose desires and beliefs have not changed, will not be rational to decide to do anything other than  $\phi$ -ing. Second, while an agent who decides to  $\phi$  will be disposed to do  $\phi$  rather than believed alternatives to  $\phi$ -ing, their volition won't dispose them not to attend to those alternatives, and won't dispose them not to act on desires for those alternatives, if such desires arise. Volitions thus do not protect against temptation. Fourth, volitions are only outputs of

---

<sup>55</sup> This conditional is surely subject to counterexamples and would need to be revised; for discussion of a related point, see, e.g., (Asarnow 2019).

<sup>56</sup> On some conceptions, volitions are as stable as intentions; compare (McCann, 1991).

<sup>57</sup> The introduction of volitions thus allows for an implementation of a decision-theoretic norm of rationality that is consistent with mentalism, in Conee and Feldman's sense. See (Conee & Feldman, 2001), as well as (Wedgwood, 2017) and (Asarnow 2019).

<sup>58</sup> Though see (Asarnow 2019) for complications.

<sup>59</sup> Are volitions a type of intention? Or are they distinct types of states? Do planning agents have volitions as well as intentions? I leave these questions open, relying only on the claim that it is possible for a creature to have a mental state with the functional role I ascribe to volitions, and that such creatures could engage in Creature 4 Sociality.

reasoning, not inputs, and so do not pose means-ends problems. However, the norm of expected desire-satisfaction maximization can de facto support a kind of means-ends coherence. When an agent very strongly desires that  $\phi$  obtain, and believes that their  $\psi$ -ing is a necessary means to  $\phi$ , if they deliberate, they will often be rational to decide to  $\psi$ .

Finally, it's worth contrasting volitions with the Searlean 'intentions-in-action' appealed to by minimalists such as Tollefsen and Blomberg.<sup>60</sup> The key role of those states is to specify the action which is performed intentionally: when someone  $\phi$ s intentionally, they had an intention-in-action to  $\phi$ . As I understand those theorists, this functional role exhausts the concept of an intention-in-action. If (as I assumed in Sections 3 and 4) belief-desire agents can act intentionally, then the desires in virtue of which their actions are intentional count as intentions-in-action. By contrast, volitions have additional features. As discussed above, They are (of necessity) products of decisions, they are subject to requirements of rationality, and they characteristically persist across short periods of time. They are like intentions-in-action in not supporting future planning, and not playing the other diachronically significant functional roles of intentions in Bratman's sense.<sup>61</sup> But, since they can be formed prior to action, they may count as prior intentions (in Searle's sense).

I claim volitions allow us to construct a creature whose social interactions will be less precarious than those of Creature 3:

Creature 4      This agent has beliefs, desires, and volitions, and has cognitive-conative theory of mind.

---

<sup>60</sup> See, e.g., (Blomberg, 2011; Searle, 1983; Tollefsen, 2005, pp. 88–89). Tollefsen's focus is on children who lack theory of mind but can represent others' behavior as goal-directed, and so can attribute intentions-in-action.

<sup>61</sup> The classic discussion of the distinction between intentions-in-action and proximal intentions (volitions) is (Mele, 1992, Chapter 10).

Crucially, just as planning agents can sometimes intend that another agent act in a certain way, agents with volitions can decide that another agent will act in a certain way.<sup>62</sup>

So equipped, Creature 4 can manifest the following type of sociality.

#### Creature 4 Sociality

- (1) Each creature truly believes that both creatures are J-ing.
- (2) Each creature decides that (a) both creatures J (b) in part by way of their volitions to J, and (c) that each Js if and only if the other Js in part by way of the other creature's volition to J.
- (3) Each creature decides that their continued J-ing will be responsive to the volitions and actions of the other.
- (4) Each creature decides that each will not do anything incompatible with what the other agent decides.
- (5) That (1)-(4) obtain is common true belief.

Success cases of Creature 4 Sociality are less precarious than Creature 3 Sociality in both ways. Because of the durable link between volition and action, it is no accident when Creature 4 agents act on their volitions. They will often engage in the kind of means-ends reasoning described above, which can lead them to engage in helping behavior, as well as bargaining and shared decision-making. So their motivations are significantly durable.

Moreover, a Creature 4 agent's volitions implicate them in a web of requirements of rationality that support their shared action in a way analogous to a planning agent's. Given persistence of underlying desires, a Creature 4 agent will be irrational to decide not to J if it reopens deliberation, and will typically be irrational not to decide to take believed necessary means to the other creature's J-ing, if they deliberate about it. Arguably, such agents are also rationally required to act on their volitions, as well. And the Creature 4 agents' acceptance of these norms further supports the durability of their dispositions.

---

<sup>62</sup> Perhaps subject to the settle condition; see section 2.2 above.

The precariousness objection thus fails to refute the parity thesis. I am skeptical of the claim that Creature 3 Sociality is more precarious (and so less thoroughly social) than Planning Agent Sociality. But even if I allowed that claim, Creature 4 Sociality shows that suitably non-precarious social interactions are possible without intentions, in Bratman's sense. A final revision of the parity thesis would thus state that if cases of spontaneous, temporary Planning Agent Sociality count as 'shared agency,' so do success cases of Creature 4 Sociality.<sup>63</sup>

## 7. Conclusion

I draw three lessons from this discussion. The first is simply the parity thesis: if planning agents engaged in spontaneous, temporary walking together count as engaging in shared agency, so do non-planning agents. The functional differences between the psychological states underlying their behavior are insignificant in this context. It follows that there can be shared agency without shared intention, and the strong form of the primacy of intention thesis is false.

The second lesson is the corollary of the parity thesis: socially extended agency does not require temporally extended agency. This means that theorists of more-than-minimal shared agency should focus not only on intention, but also on the other psychological features that can underlie the capacity for such shared agency. On my account, these include theory of mind, the taste for sociality, and perhaps being subject to requirements of rationality.

The final lesson is a product of the method of creature construction: focusing in a fine-grained way on psychological capacities other than intention has the potential to illuminate important debates in social ontology concerning the minds of creatures who do not engage in

---

<sup>63</sup> One objection here is worth noting. Bratman and others have understood an agent's acceptance of requirements of rationality in terms of intentions to conform to the requirements (at, e.g., (Bratman, 2014, pp. 16–17)). If norm acceptance necessarily involves such intentions, then Creature 4 agents must have intentions. But I doubt norm acceptance necessarily involves intentions, since such acceptance does not require the hierarchical planning capacities intentions make possible. For other theories of acceptance see, e.g., (Railton, 2006; Sripada & Stich, 2006).

shared agency. To pick one important example, Michael Tomasello and co-authors have influentially argued that the great apes do not engage in more-than-minimal shared agency, though they may be planning agents and (at least) they have enough theory of mind to distinguish intentional actions from errors and mistakes.<sup>64</sup> This thus raises the question of which psychological capacities great apes lack, in virtue of which they do not engage in shared agency.<sup>65</sup> But similar questions can be raised about other kinds of creatures.

My creature construction highlights two psychological features that can contribute to the capacity for shared agency that have often been overlooked, which one might hypothesize that great apes or other non-socially extended agents lack. The first is the taste for sociality: it may be that the great apes are not motivated to act together with other agents. The taste for sociality involves two components: the capacity to have desires or intentions that range over the behavior of other agents, and actually having desires that other agents act in parallel with oneself. Both are cognitive achievements whose significance should not be overlooked. They differ from merely affective responses to the behavior of others (such as liking or disliking it) in that they can lead to the forms of helping behavior characteristic of shared agency. One might hypothesize about agents who do not engage in shared agency that they lack the taste for sociality.<sup>66</sup>

The second candidate is cognitive-conative theory of mind. Substantial evidence suggests that great apes can distinguish intentional actions from mistakes. This has sometimes been taken as evidence that they have what I call cognitive theory of mind: they can form representations (of

---

<sup>64</sup> See, e.g., (Tomasello, 2009; Tomasello & Carpenter, 2007; Tomasello et al., 2005)

<sup>65</sup> Bratman asks this question at (Bratman, 2014, p. 159,183).

<sup>66</sup> Tomasello and co-authors briefly suggest a similar hypothesis at (Tomasello et al., 2005, p. 686). That being said, Tomasello, with Felix Warneken, has also presented evidence that chimpanzees engage in helping behavior similar to that characteristic of shared agency, which might suggest they have desires about the behavior of others. See, e.g., (Warneken & Tomasello, 2009). More recently, other authors have suggested the relevant behavior is an experimental artifact, and that chimpanzees do not help in this way in general; see (Tennie, Jensen, & Call, 2016).

some kind) about the mental states of others.<sup>67</sup> My creature construction, however, highlights the importance of cognitive-conative theory of mind: of the ability to form desires or volitions about the desires and volitions of other creatures. So it suggests the hypothesis that some creatures which lack the capacity for shared agency lack it in virtue of lacking cognitive-conative theory of mind.

I do not mean to overstate the importance of this final lesson, and of course I present no evidence concerning whether either of these hypotheses is plausible in the case of the great apes. The broader point is research in social ontology will benefit from paying close, fine-grained attention to the mental capacities other than intention that can underlie the capacity for shared agency.<sup>68</sup>

---

<sup>67</sup> Great ape theory of mind is highly controversial; compare (Call & Tomasello, 2008) and (Martin & Santos, 2016).

<sup>68</sup> For helpful conversations about these topics, or written comments on drafts, I thank Facundo Alonso, Michael Bratman, Sarah Buss, Henry Kellison, Carlos Núñez, Matthew Rachar, Grant Rozeboom, and two anonymous referees.

## References

- Alonso, F. M. (2016) 'A Dual Aspect Theory of Shared Intention,' *Journal of Social Ontology*, 2/2, 271–302.
- Alonso, F. M. (2018) 'Reductive Views of Shared Intention,' In M. Jankovic & K. Ludwig (Eds.), *The Routledge Handbook of Collective Intentionality*. New York: Routledge.
- Andrews, K. (2016) 'Animal Cognition.' In *Stanford Encyclopedia of Philosophy*. Retrieved from <https://plato.stanford.edu/entries/cognition-animal>
- Asarnow, S. (2019) 'On Not Getting Out of Bed,' *Philosophical Studies* 176/6, 1639–66
- Baker, D. (2010). 'Ambivalent Desires and the Problem with Reduction,' *Philosophical Studies*, 150/1, 37–47.
- Baker, J. (2008). 'Rationality without Reasons,' *Mind*, 117/468, 763–782.
- Blomberg, O. (2011) 'Socially Extended Intentions-in-Action,' *Review of Philosophy and Psychology*, 2/2, 335–353.
- Blomberg, O. (2015a) 'An Account of Boeschian Cooperative Behaviour,' In C. Misselhorn (Ed.), *Collective Agency and Cooperation in Natural and Artificial Systems: Explanation, Implementation and Simulation*. Cham: Springer.
- Blomberg, O. (2015b) 'Shared Goals and Development,' *The Philosophical Quarterly*, 65/258.
- Blomberg, O. (2016) 'Common Knowledge and Reductionism about Shared Agency,' *Australasian Journal of Philosophy*, 94/2.
- Bratman, M. E. (1987). *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press.
- Bratman, M. E. (1992) 'Shared Cooperative Activity,' *The Philosophical Review*, 101/2, 327–341.
- Bratman, M. E. (2006) 'Dynamics of Sociality,' *Midwest Studies In Philosophy*, 30, 1–15.
- Bratman, M. E. (2007) 'Valuing and the Will,' In *Structures of Agency*. New York: Oxford University Press.
- Bratman, M. E. (2014) *Shared Agency*. Oxford: Oxford University Press.
- Bratman, M. E. (2018a) 'Intention, Belief, and Instrumental Rationality,' In *Planning, Time, and Self-Governance*. Oxford: Oxford University Press.
- Bratman, M. E. (2018b) 'Introduction: The Planning Framework,' In *Planning, Time, and Self-Governance*. Oxford: Oxford University Press.
- Bratman, M. E. (2018c) *Planning, Time, and Self-Governance*, New York: Oxford University Press.
- Butterfill, S. A. (2012) 'Joint Action and Development,' *The Philosophical Quarterly*, 62/246, 23–47.
- Butterfill, S. A. (2015) 'Planning for Collective Agency,' In C. Misselhorn (Ed.), *Collective Agency and Cooperation in Natural and Artificial Systems: Explanation, Implementation and Simulation*. Cham: Springer.
- Butterfill, S. A. (2016) 'Joint action: A minimalist approach,' In J. Kiverstein (Ed.), *The Routledge Handbook of Philosophy of the Social Mind*.
- Butterfill, S. A., & Sebanz, N. (2011) 'Editorial: Joint Action: What Is Shared?' *Review of Philosophy and Psychology*, 2/2, 137–146.
- Call, J., & Tomasello, M. (2008) 'Does the chimpanzee have a theory of mind? 30 years later,' *Trends in Cognitive Sciences*, 12/5, 187–192.



- Chant, S. R., & Ernst, Z. (2007) 'Group intentions as equilibria' *Philosophical Studies*, 133/1, 95–109.
- Conee, E., & Feldman, R. (2001) 'Internalism Defended,' *American Philosophical Quarterly*, 38/1, 1–18.
- Davidson, D. (1980) 'Actions, Reasons, and Causes,' In *Essays on Actions and Events*. Oxford: Oxford University Press.
- Ernst, Z., & Chant, S. R. (2007) 'Collective action as individual choice,' *Studia Logica*, 86/3, 415–434.
- Gilbert, M. (1990) 'Walking Together: A Paradigmatic Social Phenomenon,' *Midwest Studies in Philosophy* 15, 1-14
- Gilbert, M. (2009) 'Shared Intention and Personal Intentions,' *Philosophical Studies*, 144/1, 167–187.
- Godfrey-Smith, P. (2005) 'Folk Psychology as a Model,' *Philosophers Imprint*, 5/6, 1–16.
- Grice, H. P. (1974) 'Method in Philosophical Psychology (From the Banal to the Bizarre),' *Proceedings and Addresses of the American Philosophical Association*, 48, 23–53.
- Kutz, C. (2000) 'Acting Together,' *Philosophy and Phenomenological Research*, 61/1, 1–31.
- Martin, A., & Santos, L. R. (2016) 'What Cognitive Representations Support Primate Theory of Mind?' *Trends in Cognitive Sciences*, 20/5, 375–382.
- McCann, H. J. (1991) 'Settled Objectives and Rational Constraints,' *American Philosophical Quarterly*, 28/1, 25–36.
- McMahon, C. (2005) 'Shared Agency and Rational Cooperation,' *Noûs*, 39/2, 284–308.
- Mele, A. R. (1992) *Springs of Action*. Oxford: Oxford University Press.
- Miller, S. (2004) *Social Action: A Teleological Account*, Cambridge: Cambridge University Press.
- Núñez, C. (2019) 'Requirements of intention in light of belief,' *Philosophical Studies*.
- Pacherie, E. (2007) 'Is collective intentionality really primitive?' *Mental Processes: Representing and Inferring*, 62/1, 153–175.
- Pacherie, E. (2011) 'Framing Joint Action,' *Review of Philosophy and Psychology*, 2/2, 173–192.
- Pacherie, E. (2013) 'Intentional joint agency: shared intention lite' *Synthese*, 190, 1817–39.
- Paternotte, C. (2015) 'Constraints on Joint Actio,' In J. Michael & M. Gallotti (Eds.), *Perspectives on Social Ontology and Social Cognition*. Dordrecht: Springer.
- Railton, P. (2006) 'Normative Guidance,' In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics* (Vol. 1).
- Railton, P. (2014) 'Reliance, Trust, and Belief,' *Inquiry*, 57/1, 122–150.
- Searle, J. (1983) *Intentionality*. Cambridge: Cambridge University Press.
- Searle, J. (1990) 'Collective Intentions and Actions,' In P. R. Cohen, M. E. Pollack, & J. Morgan (Eds.), *Intentions in Communication*. Cambridge, MA: MIT Press.
- Sober, E., & Wilson, D. S. (1998) *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press
- Sripada, C. S., & Stich, S. (2006) 'A Framework for the Psychology of Norms,' In P. Carruthers, S. Laurence, & S. Stich (Eds.), *Innate Mind: Culture and Cognition* (Vol. 2). Oxford: Oxford University Press.
- Tennie, C., Jensen, K., & Call, J. (2016) 'The nature of prosociality in chimpanzees,' *Nature Communications*, 7, 1–8.
- Tollefsen, D. (2005) 'Let's pretend! Children and joint action,' *Philosophy of the Social Sciences*, 35/1, 75–97.

- Tomasello, M. (2009). *Why We Cooperate*. Cambridge, MA: MIT Press.
- Tomasello, M., & Carpenter, M. (2007) 'Shared intentionality,' *Developmental Science*, 10(1), 121–125.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005) 'Understanding and sharing intentions: The ontogeny and phylogeny of cultural cognition,' *Behavioral and Brain Sciences*, 28, 675 – 691.
- Tuomela, R. (2005) 'We-Intentions Revisited,' *Philosophical Studies*, 125/3, 327–369.
- Tuomela, R. (2009) 'Collective Intentions and Game Theory,' *The Journal of Philosophy*, 106/5, 292–300.
- Velleman, J. D. (2000) 'How To Share An Intention,' In *The Possibility of Practical Reason*. Oxford: Oxford University Press.
- Vesper, C., Butterfill, S. A., Knoblich, G., & Sebanz, N. (2010) 'A minimal architecture for joint action,' *Neural Networks*, 23/8–9, 998–1003.
- Warneken, F., & Tomasello, M. (2009) 'The roots of human altruism,' *British Journal of Psychology*, 100/3, 455–471.
- Wedgwood, R. (2017). *The Value of Rationality*. Oxford: Oxford University Press.