

ORIGINAL ARTICLE

How to be minimalist about shared agency

Jules Salomone-Sehr 

The Queen's College, University of
Oxford, Oxford, UK

Correspondence

Jules Salomone-Sehr, The Queen's
College, University of Oxford, Oxford,
United Kingdom.

Email:

jules.salomone-sehr@queens.ox.ac.uk

Funding information

Canada Research Chair on Epistemic
Injustice and Agency

Abstract

What is involved in acting together with others? Most shared agency theorists endorse the Shared Intention Thesis, i.e., the claim that shared agency necessarily involves shared intentions. This article dissents from this orthodoxy and offers a minimalist account of shared agency—one where parties to shared activities need not form rich webs of interrelated psychological states. My account has two main components: a conceptual analysis of shared agency in terms of the notion of plan, and an explanation of undertheorized agency-sharing mechanisms. My analysis states that we act together just in case our activities conform to a plan and that plan figures in an explanation of our activities' joint conformity to it. To sloganize: *shared activity is plan-coordinated activity*. Sometimes, plan-coordination goes by way of shared intentions. However, besides shared intentions, there are at least two additional families of agency-sharing mechanisms. The first features a central planner who determines the content of a plan and attributes the different parts of that plan to a collection of agents. The second does away with the planner and involves a roughly Darwinian selection of patterns of activity. Both families of mechanisms enable us to act together even in the absence of shared intentions.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Philosophy and Phenomenological Research* published by Wiley Periodicals LLC on behalf of Philosophy and Phenomenological Research Inc.

KEYWORDS

coordination, minimalism, plan, shared activity, shared agency, shared intention

1 | INTRODUCTION

Many of the things we do are the expression of our individual agency: turning off the alarm, drinking one's morning coffee. At times these things we individually do trigger sequences of events that snowball into non-additive outcomes: my heavy coffee consumption, combined with the similarly heavy consumption of many coffee drinkers, might put some strain on water resources in regions where coffee is produced and, thus, disturb these regions' ecosystems. But while individual actions might sometimes amass to cause such large-scale effects, they do not thereby together count as shared activity. In addition to engaging in individual actions (that might bring about non-additive outcomes), we also often engage in *shared activities*, that is, in activities done genuinely together. Shared agency, the topic of this article, is fundamental to our rich social lives. We make movies, dance, rob banks, attend protests, and sit in Parliament together. Besides, only shared efforts are likely to solve the urgent challenges we face today as individuals, communities—and, it turns out, as a species.

In this article, I offer a *minimalist* account of shared agency. Specifically, I argue that our activities are shared just when they are coordinated by a common plan of action that might, but need not, figure in the content of shared intentions.

2 | THE NEED FOR A MINIMALIST ACCOUNT

Why do we need a minimalist account of shared agency?

Shared agency theorists, no less than individual agency theorists, have shown a bias towards the conscious and the intentional—to use David Velleman's phrase, a bias towards action *par excellence* (2015). In shared agency theory, this bias has found its most significant manifestation in the widely held claim that shared intentions are *necessary* for the performance of shared activities (hereafter, the *Shared Intention Thesis*).¹ Prominent accounts of shared agency have certainly met some success in their treatment of shared activities where participants, in small number, are tightly knitted and their psychological states narrowly overlap (as when we paint our house together). By contrast, they have had a much harder time handling cases where participation is automatic and unthinking (as when we get off a packed NYC subway),² reluctant (as when the

¹ Here is a non-exhaustive list of shared agency theorists who have committed themselves to some version or other of the Shared Intention Thesis: John Searle (1990, p. 402), Christopher Kutz (2000, p. 74–81), Abraham Roth (2004, p. 361), Philip Pettit and David Schweikard (2006, p. 23–24), and Facundo Alonso (2009, p. 445). Here are some exceptions: Sara Rachel Chant (2006), Stephen Butterfill (2012, 2016), Scott Shapiro (2014), Katherine Ritchie (2020), Samuel Asarnow (2020), and Jules Salomone-Sehr (2022, 2023). Two of the most influential accounts of shared agency, namely, those of Margaret Gilbert and Michael Bratman, are largely consistent with the Shared Intention Thesis. However, both Gilbert (2013, p. 12) and Bratman (2014, p. 169, endnote 86) seem open to the idea that there are other routes to shared activity than shared intentions.

² See Elisabeth Pacherie (2013) for an attempt to handle relatively unthinking shared activities while avoiding open conflict with the Shared Intention Thesis.

firm that we have little choice but to work for pursues dubious ends),³ and/or where the number of participants is high (as when we get a shot as part of a vaccination campaign).⁴ What is more, shared agency theorists have too quickly assumed that unknowing participation in shared agency is a conceptual impossibility.⁵ The bias towards agency par excellence, in short, has led action theorists to an excessively narrow understanding of shared agency.

The problem with this narrow understanding of shared agency is not only that it might misclassify some activities done together as mere parallel individual actions—a problem of extensional inadequacy. This narrow understanding has also failed to do justice to the variety of jobs that the concept of shared agency does for us.

Prominent accounts have primarily focused on the perspective of voluntary participants in shared activities. For such participants, understanding their intentional conduct as a contribution to a shared activity surely primes dispositions for mutual help, mutual responsiveness, and/or team thinking—that is, the sort of dispositions that shared agency theorists have sometimes included in their accounts of shared agency (Bratman, 1992, 2014; Gold & Sugden, 2007; Rachar, 2023).

However, priming such cooperative and phenomenologically rich dispositions is not the only job that the concept of shared agency does for us. For one thing, the possession of this concept enables us to *plan* shared activities, and from the perspective of the planners of such activities, it may not matter much whether every participant intends the achievement of some common goal and/or the enactment of some common plan. What certainly matters instead is that through shared agency, one way or another, things get done—an outcome made sometimes more likely if the shared performance is suitably insulated from the vicissitudes of participants' psychologies.

The concept of shared agency is also usefully deployed in social critique of oppressive patterns of collective activity. For instance, a critique of gender roles might observe that when a woman in a straight relationship does the majority of domestic chores, she is not acting merely in her capacity as individual agent; instead, she has been sharing all along her agency with her husband as part of a plan (possibly not of her own making, of course) in which he gets to have a career and she ends up in the backseat, taking care of the logistics of his success. In examples of this sort, one aim of the social critique is to reveal that what might have seemed to be expressions of individual agency are in fact manifestations of a common, yet oppressive, plan—one in the enactment of which women might even feel complicit.⁶ Another is to shed light on the mechanisms that explain the persistence of that common plan, and in particular the oppressed party's ongoing, and perhaps unknowing, participation in its enactment.⁷

³ Kutz (2000) and Bratman (2014), two Shared Intention Thesis sympathizers, have both attempted to accommodate reluctant participation.

⁴ For an instance of attempt to reconcile the Shared Intention Thesis with shared activities involving many participants, see Cédric Paternotte (2014).

⁵ Kirk Ludwig is an exception. His account leaves indeed plenty of room for unwitting participation in shared agency (2016, p. 219–221). See also Christian List and Philip Pettit (2011, p. 32–33).

⁶ Whether they are right to feel so is a thorny question that it is not this article's job to explore. Given my present purposes, suffice it to say that understanding the enactment of oppressive plans to be something done *together* explains why wondering about the complicity of the oppressor as well as that of the oppressed makes sense. On an everyday conception of complicity, indeed, complicitous responsibility is responsibility that one bears for one's participation in shared wrongdoing, and hence seems to involve some sharing of agency—an intuitive truth captured by Kutz's account of complicity (2000).

⁷ See, e.g., Serene Khader (2019, p. 72–73), Cailin O'Connor (2019, p. 3–4), Charlotte Knowles (2019), and Manon Garcia (2021).

Shared agency, thus, is not just a useful label to refer to endeavors that we join in consciously and intentionally (i.e., the sort of endeavors that shared agency theorists have predominantly focused on). Another of the functions of the concept of shared agency is to help us navigate our social world by recognizing the endeavors that more or less unintentionally, and perhaps even unknowingly, we might have carried out together.

For all these reasons, we need a *minimalist* account of shared agency, one where the participants' involvement does not crucially depend on their intentions and doxastic states—one, then, that rejects shared agency theorists' bias towards action par excellence and their related commitment to the Shared Intention Thesis.⁸

3 | KEEPING THE MINIMALISM IN CHECK: THREE DESIDERATA

I have just explained what motivates my search for a minimalist account. This minimalism, however, should not overshoot; otherwise, my account would fail to capture a distinctive phenomenon of shared agency. To keep this sought-for minimalism in check, I now offer three adequacy criteria that accounts of shared agency must meet.

First, adequate accounts of shared agency should not conflate multi-agent causation and shared activity. There is, indeed, no need to invoke the notion of shared agency to explain cases where individual agents, *acting merely in parallel*, cause some outcome or other. In such cases, individual agency does the explanatory trick just fine.⁹

Second, adequate theories of shared agency must account for the fact that if φ is a shared activity, then there must be some common practical orientation that its constituent individual activities all follow.¹⁰ An individual activity is not a mere jumble of disconnected bodily movements; rather, it evinces some unified practical orientation. Similarly, a shared activity is not a mere jumble of disconnected individual activities: it, too, evinces some unified practical orientation. Put differently, our activities are not shared unless there is some shared specification of how our activities are all supposed to go.

Third, adequate theories of shared agency should not allow, when φ is a shared activity, that its constituent individual activities follow a common practical orientation by mere happenstance.¹¹ To see this, imagine that, on a sunny day, you and I are cleaning up Malibu beach by picking plastic bottles. Our picking of plastic bottles arguably follows a common practical orientation: after all, what each of us is doing aims at the same end, namely, that Malibu beach be clean. We would not be engaged in shared agency, however, if you and I coincidentally happened to be cleaning

⁸ See Kutz (2000, p. 66–112), Butterfill (2012, 2016), and Butterfill and Sinigaglia (2023) for other minimalist accounts. Because he endorses the Shared Intention Thesis, Kutz's account is, I believe, insufficiently minimalist. See endnote 17 below for further elaboration. I discuss Butterfill and Sinigaglia at the end of §6.

⁹ Chant's account does not meet this desideratum, for she claims that behaviors that cause non-additive outcomes (i.e., outcomes for the production of which individual contributions are non-separable) count as collective actions (2006, p. 430). (To be fair, it might be that Chant's account of collective action is not meant to elucidate the phenomenon of shared agency that is the focus of this article.)

¹⁰ This desideratum echoes Paul Grice's observation that cooperative efforts always involve a "mutually accepted direction" (1989, p. 26).

¹¹ For a similar view, see Kutz (2000, p. 76).

up Malibu beach on the same day: despite aiming at the same end, each of our picking of plastic bottles would involve no more than individual agency.¹²

To summarize this discussion, even on a minimalist account of shared agency such as the one I am set to offer in this article, the sense in which φ is done together, when φ is a shared activity, is fairly *robust*: it is the ‘together’ of a common practical orientation that constituent individual activities of φ all follow non-coincidentally.

4 | A PUZZLE

It might seem, initially, that the most promising way to arrive at a plausible account of shared agency that satisfies the desiderata I just spelt out is to endorse the Shared Intention Thesis.

First, activities that satisfy intentions we share are obviously excellent candidate examples of shared activities, and not mere examples of multi-agent causation.

Second, the claim that shared intentions are necessary for shared agency explains effortlessly the common practical orientation shown by shared activities: the content of such shared intentions specifies how the participants’ constituent actions are supposed to go.

Third, activities that satisfy intentions we share follow a common practical orientation non-coincidentally. When enacting a shared intention, our constituent activities do not conform to our shared intention’s content by mere happenstance. Rather, our shared intention explains our individual activities’ joint conformity to that shared intention’s content. That is to say: our shared intention explains not just why you act in conformity with some common goal and/or plan and why I act in conformity with that same goal or plan, but also why your conformity and mine go hand in hand. At least that much is the case under prominent accounts of shared intentions, especially those that include a condition of mutual responsiveness (e.g., Bratman, 2014).¹³

To recap, the Shared Intention Thesis paves the way to an account of shared agency that satisfies all three desiderata.

The Shared Intention Thesis, however, clashes with the minimalism that I am after (§1-2). Furthermore, despite its initial plausibility, the Shared Intention Thesis is false.

To see this, consider the dancers of a ballet company. I contend that they do not have to share the intention to execute their choreography, or indeed any other intention, to execute the

¹² True, we might be said to be cleaning up the beach *together*. The relevant sense of ‘together’, however, is merely causal: yours and my picking up of plastic bottles causally contribute to a cleaner beach.

¹³ In the literature, there are perhaps accounts of shared intentions and agency on which we might count as enacting our shared intention (and hence as acting together) even when the fact of our acting in conformity to the goal or plan that figures in that shared intention’s content is merely coincidental. Ludwig’s account might be one such account (2016). Imagine that you and I, in the absence of any coordination, have each randomly formed the intention that we both jump at the same time in accordance with a shared plan. On Ludwig’s account, we count as sharing an intention, and enacting that shared intention is, on his view, just a matter of each of us enacting our respective intentions that we jump simultaneously in accordance with a shared plan. Should we enact our respective intentions, we would therefore count, on Ludwig’s account, as jointly intentionally jumping together. But this is odd, given that acting together takes more than a mere appearance of coordination. On the surface, Ludwig seems to agree with this point insofar as he contends that genuine coordination (even if minimal) is essential to shared agency (2016, p. 225). This contention, however, is hard to square with the fact that his account classifies the random jumping scenario as an instance of shared agency even though our acting in conformity to the goal of jumping at the same time involves no coordination. (I owe this criticism of Ludwig’s account as well as the random jumping scenario to an anonymous referee.)

choreography together: closely sticking to their individual choreographed roles does the trick just fine.

Let me flesh this out. Suppose you are a narcissistic dancer of that ballet company and wish that you did not have to share the audience's attention with the other dancers. Should we fail to show up for the performance, you would not bother to contact us and ask us what happened; instead, you would gladly perform alone. If that were the case, you would not be committed to the collective nature of our performance. All the same, even if, just like you, we all stuck narrowly to our respective parts, we would still be giving a performance together (provided of course that our conformity with the choreography is non-coincidental, for instance because our dance moves were designed by a choreographer who distributed the different parts of the choreography to all of us).¹⁴ This argument, inspired by Scott Shapiro (2014),¹⁵ suggests that shared intentions are not necessary for shared agency.

There is a complication (Salomone-Sehr, 2022). A collection of individuals might be said to share intentions either distributively or collectively. Roughly:

- An intention is shared by some collection of individuals in the *collective* sense just when the collection itself has that intention.
- An intention is shared by some collection of individuals in the *distributive* sense only if, at the very least, every individual of that collection has an intention whose content narrowly overlaps with the content of the others' intentions.¹⁶

The argument developed above shows that intentions shared in the distributive sense are not necessary for shared agency: when we stick to our respective roles, my intention and yours do not have narrowly overlapping contents. And yet, we are clearly engaged in shared agency. But what about intentions shared in the collective sense? Might they be required for shared agency even if intentions shared in the distributive sense are not?

No. To see this, suppose that the ballet company's performance has been designed by an authoritarian choreographer who calls all the shots. Suppose, additionally, that every dancer's move is solely motivated by their individual intention to stick to their respective role in the choreography—thus, suppose that they share no intention in the distributive sense. In this scenario, there is no point in attributing the intention to enact the choreography to the entire dance company, taken together: this intention is merely the authoritarian choreographer's own. There is, therefore, no intention that the dancers of this scenario can be truly said to share, either distributively or collectively. And yet, the dancers are performing together.¹⁷

¹⁴ This anticipates §7.1 below.

¹⁵ See also Ritchie (2020).

¹⁶ In addition to the narrow overlap of individual intentions' contents, advocates of the Shared Intention Thesis who take a distributive approach have often included other conditions in their accounts of the sharing of intentions (e.g., mutual responsiveness, common knowledge, etc.). That is why I did not say 'if and only if' above.

¹⁷ See Salomone-Sehr (2022). Notice that the conditions that, as laid out above, must be satisfied by a collection of agents for them to count as sharing intentions (either distributively or collectively) are rather *minimal* conditions; that is, they are conditions that *even on a minimalist account of shared intentions*, a collection of agents must satisfy to count as having a shared intention. This is an important point, for it shows that an advocate of the Shared Intention Thesis who goes for a minimalist account of shared intentions will be forced to conclude, mistakenly, that the narcissistic dancers working under the authoritarian choreographer (as in the case developed above) are not engaged in shared agency. To handle this case adequately, therefore, we need a *genuinely* minimalist account of shared agency, one on which shared intentions,

This denial of the Shared Intention Thesis and the desiderata I argued for earlier create a puzzle. If shared intentions are not necessary for shared agency, then they do not always supply the common practical orientation evinced by shared activities. When they do not, *what will?*

5 | APPROACH

I am now about to present my positive account, namely a minimalist account that satisfies the above adequacy criteria (§3) while doing away with the Shared Intention Thesis (§4).

But first, let me specify my approach.

5.1 | Keeping conceptual analysis and mechanism explanation separate

My view has two main components: an analysis of the concept of shared agency in terms of the notion of plan of action, and an account of the mechanisms that enable the formation and enactment of plans, and therefore that provide for shared agency. Proponents of the Shared Intention Thesis have not made this distinction. In fact, the Shared Intention Thesis combines both *mechanism explanation* and *conceptual analysis*: after all, the enactment of shared intentions is an agency-sharing mechanism—that is, a mechanism thanks to which agency is shared. But as argued earlier, shared intentions are not necessary for shared agency.

even on a minimalist conception thereof, are not necessary. To illustrate this point, consider Kutz's minimalist account of shared intentions. His account is minimalist in part because, contrary to most accounts, it does not require that sharers of the intention to φ intend the achievement of φ (2000, p. 100). It merely requires, instead, that they all have the *participatory intention* to φ —that is the intention to contribute instrumentally, constitutively, or expressively to φ (p. 82). (In that, Kutz's account satisfies the condition of narrow overlap stated above in the main text.) Now, even on Kutz's minimalist account, it seems odd to describe the narcissistic dancers as sharing the intention to perform in a tableau: rather than intending to contribute to a tableau, they each merely intend to stick narrowly to their roles in the choreography—*while not intending, thus, that their roles form part of the choreographed tableau*. In fact, should it turn out that, for some dancer, sticking to their role no longer amounts to a contribution to a tableau (say, because the choreographer made the decision to get rid of all the other dancers and, so, chose to transform the tableau into a solo), that dancer would not alter their conduct so as to try and contribute to a tableau (something that, according to Kutz [p. 101], they would have had to do, had they had the participatory intention to perform in that tableau). Rather, that dancer would gladly dance alone—and receive undivided attention from the audience. Even on Kutz's minimalist account of shared intentions, therefore, the narcissistic dancers do not share the intention to perform in a tableau. Now, insofar as Kutz regards overlapping participatory intentions as the core of shared agency, it follows in turn that, on his account, the narcissistic dancers are *not* engaged in shared agency. But this is a mistake, as we saw earlier. (To be sure, Kutz argues at some point [p. 103] that participatory intentions to φ are not even necessary for φ to count as a shared activity and that knowing contributions to φ might suffice. However, it is important to stress, first, that this claim is in evident tension with Kutz's official view—i.e., the view that, at its core, shared agency requires nothing more *and nothing less* than overlapping participatory intentions [e.g., p. 75]. In addition, it bears noting that an account on which knowing contributions to φ suffice for shared agency opens the floodgates of overgeneration. For instance, such an account would classify heavy coffee drinkers who knowingly contribute to depleting water resources in regions where coffee is produced as engaged in shared agency when so contributing. In this example, however, depleting water resources should strike us as the outcome of nothing more than multi-agent causation.) In passing, a consequence of the foregoing discussion is that Kutz's account of shared agency differs extensionally from the minimalist account that later sections of my article develop. I thank an anonymous referee for pressing me to clarify how my views relate to Kutz's.

There must then be other agency-sharing mechanisms besides shared intentions, and it is therefore a mistake to analyze the concept of shared agency in terms of the concept of shared intentions.

Here is what to expect in what follows: I first offer a mechanism-neutral analysis of shared agency in terms of the notion of plan (§6), and then describe two mechanisms that provide for shared agency even in the absence of shared intentions (§7). In so doing, I keep conceptual analysis and mechanism explanation separate.

5.2 | Necessary and sufficient conditions

Keeping conceptual analysis and mechanism explanation separate in shared agency theory is also something that Stephen Butterfill and Corrado Sinigaglia do. They have indeed recently argued that an illuminating conceptual elucidation of shared agency can remain neutral about the mechanisms that provide for shared agency (2023). Despite their similarities, my approach is, in one respect, less open-ended than theirs. Like them, I believe that there is not just one, but a plurality of mechanisms that provide for shared agency (as I argue in §7). By contrast, I am less tempted by their idea that there might be not just one, but a plurality of analyses of shared agency that are all equally illuminating. In the next section (§6), I will thus aim to provide what I take to be the one true analysis of shared agency (at least, that's the hope), an analysis that I will accordingly articulate in terms of not just sufficient, but also *necessary* conditions.

There are two metatheoretical reasons why I aim to couch my analysis in the rather traditional format of necessary and sufficient conditions. One is that it would be desirable, at least on the grounds of theoretical parsimony, if the plurality of agency-sharing mechanisms was nevertheless compatible with a single, unified characterization of shared agency—the kind of characterization that necessary and sufficient conditions are particularly well positioned to articulate. The resulting picture would be a form of ordered pluralism,¹⁸ one on which we might think of the different agency-sharing mechanisms as different species of a single genus, namely, the genus of shared agency. In that respect, the characterization of shared agency I defend in the next section might be regarded as a *generic* characterization.¹⁹

There is a second reason why I aim not just for sufficient, but also for *necessary* conditions. As said at the outset of the article, the account of shared agency I offer here is a *minimalist* account, one on which shared intentions are not necessary for shared agency. Now, conceptual elucidations articulated in terms of necessary (and sufficient) conditions are biased in favor of minimalism: for any feature initially thought to be necessary for the characterization of one's target phenomenon, it takes indeed just one instance of that target phenomenon where that feature is missing to conclude that the feature in question, for failing to be necessary, deserves to be removed from one's account. This bias in favor of minimalism might sometimes be unwelcome. One's target phenomenon might be such that explanatorily illuminating features of that

¹⁸ I borrow this expression from Miranda Fricker (2019). Butterfill and Sinigaglia suggest (rightly, I believe) that their arguments delineate a sort of disciplined pluralism about shared agency (2023, p. 18). However, given their openness to the idea that there might be different, equally correct conceptions of shared agency, their disciplined pluralism is less disciplined than mine.

¹⁹ Michael Bratman made that suggestion to me in conversation.

phenomenon are not found in every instance of it. For phenomena of that sort, necessary and sufficient conditions would deliver an explanatorily impoverished characterization.²⁰ But the fact that necessary and sufficient conditions are biased in favor of minimalism is not a problem for the account of shared agency that I am after—quite the opposite. As I have been arguing, we have indeed good reason to think that a minimalist account of shared agency is exactly what we need, hence the appeal of the minimalism-friendly format of necessary and sufficient conditions.

Enough with this methodological interlude. Let me now turn to my analysis of shared agency and articulate its necessary and sufficient conditions.

6 | THE PLAN-COORDINATION ANALYSIS OF SHARED AGENCY

Here is my analysis of shared agency: we act together if and only if our activities conform to a common plan of action, and that plan figures in an explanation of our activities' joint conformity to it. To sloganize: *shared activity is plan-coordinated activity*.²¹

You might wonder, What is plan-coordination? And first, what is a plan?

At the highest level of generality, a plan is a description of how an item with connecting parts is supposed to operate (Graham, 2011). For instance, the plan of a watch specifies how it is supposed to work, that is, how its different parts are supposed to behave. Similarly, then, a plan of action specifies a blueprint for a collection of activities. The plan of action behind a complex production process describes how that process is supposed to function, that is, how the tasks are divided, when and how each task should be completed, by whom, and with which tools.

Plans of action do not specify how different activities are supposed to interlock just for the sake of it: they serve a purpose (synonymously, goal) or set of purposes. Factory employees and machines are supposed to interact in a certain way not just for the thrill of it, but in order to produce certain goods.

In a more formal key, then, a plan of action is a description²² of a sequence of actions that can helpfully be modelled by the following triple:

$$P = \langle A, G, h \rangle$$

Where A is a set of actions, G , a set of purposes (e.g., events, states of affairs, etc.), and h , a function on A that represents dependency relations between members of A .²³

²⁰ See Fricker (2016).

²¹ This analysis is indebted to Bratman, Ludwig, and Shapiro's insightful suggestion that common plans are central to shared agency. Unlike Ludwig, however, I do not believe that the common plan our shared activity enacts must figure in the content of an intention we share, for we might not share any intention. And unlike Shapiro, I do not believe that parties to a common plan must be committed to their roles in it (as I argue in §7.1 with the example of the footprint stickers), nor do I believe that our plan must have been intelligently designed (as I argue in §7.2).

²² Given that plans of action are *descriptions* of the proper functioning of collections of activities, they fall under the category of abstract objects with semantic content. This, however, might seem to create a difficulty: if they are abstract objects, how can plans explain our activities' conformity to them? See endnotes 45, 50, and 54.

²³ Here, I am roughly following Myles Brand (1986, p. 219). See also Joshua Shepherd (2021, p. 8-10). I should take note of some contrasts between Brand's formal notations and mine. First, surely because Brand's focus is on plan-types, the set A , in his account, is not a set of actions, but rather of action-types. My focus, by contrast, is on *plans*, and plans are

With this in mind, let us now turn our attention to my plan-coordination analysis of shared agency. This analysis involves two conditions, namely, that our activities conform to some common plan, and that that common plan figures in an explanation of our joint conformity to it. It is relatively obvious that both conditions are individually *necessary* for shared agency.

First, I do not see any plausible sense in which activities undertaken in the absence of a common plan could count as shared. Even the unscripted performance of actors of an improvisational theater company involves the initially partial²⁴ plan to give an improv performance.

Turning to the second condition, the plan must not be explanatorily idle if activities that conform to it are to count as shared, for otherwise I would have a serious overgeneration problem: after all, any collection of compossible activities by distinct agents surely conform to some abstractly conceived plan, and so would count as an instance of shared agency if it were not for the second condition of my analysis.²⁵ Now, not only does this second condition guarantee that the plan not be explanatorily idle, it also makes sure that the plan explains not just our conformity to it, but our *joint*²⁶ conformity to it.²⁷ This important feature of my analysis excludes cases of coincidental conformity. Imagine you and I randomly form separately the intention to enact the plan that we both jump at the same time, and then proceed to enact that intention.²⁸ My account does not classify the resulting coincidental alignment of individual agencies as an instance of shared agency: although the plan we enact explains your conformity with it as well as mine, there is nevertheless no plan-involving mechanism that makes it likely that your acting in conformity with that plan and my acting in conformity with it too go hand in hand. Before moving on, notice that the explanation that links the plan to our joint conformity to it must be sufficiently *tight*: the more circuitous the path from the plan to our activities, the more controversial it will be to say that, together, we have enacted it.

In addition to being individually necessary, the two conditions I have argued for so far are also jointly *sufficient* for shared agency, as I argue now.

You might think that not any plan can provide for shared agency. More specifically, you might think that a plan can only coordinate our activities, and hence enable us to act together, if it is not

descriptions of sequences of *actions*, and not of actions-types. Second, I agree with Brand that purposes form part of the individuation conditions of plans. Unlike him, however, I think they are not best modelled as functions, but rather as members of a set—namely, the set of purposes *G* that the plan is meant to serve. The rough-and-ready model offered in the main text might need refinement in order to accommodate complex plans. One first refinement consists in allowing that, besides purposes, *G* might contain *sub*-purposes, that is, purposes that early stages of a complex plan should achieve. Here is an additional refinement needed to handle contingency plans. Contingency plans are typically activated when some purpose, or sub-purpose, has not been achieved adequately. To capture this sort of dependency relation, *h* should be defined not just on *A*, but on *A* together with *G*. The formal notations introduced in the main text can surely lend themselves to further refinements—something however that, given my present purposes, I will not explore.

²⁴ This language exploits one property of plans, namely their being susceptible of varying degrees of specificity (Brand, 1986, p. 226; Bratman, 1987). In fact, plans might sometimes be so partial that they only prescribe the pursuit of a goal (or set of goals) while remaining silent about the steps (i.e., the actions arranged according to certain dependency relations) that should be taken as part of that pursuit.

²⁵ I thank Kirk Ludwig for drawing my attention to this issue.

²⁶ I am very grateful to an anonymous referee for prompting this important clarification.

²⁷ One might worry that talk of joint conformity makes the analysis circular. This worry would be on point if “joint conformity” meant “conformity achieved by way of shared agency.” But obviously, this is not what I mean. What I mean, rather, is that the plan must explain not just each of the relevant agent’s acting in conformity with the plan, but also why their acting in conformity to that plan co-occurs.

²⁸ See endnote 13 above for discussion of a case of that sort.

defective. As my earlier characterization of plans suggests, plans might be defective in at least two ways.²⁹ Plans might be *internally* ineffective when the dependency relations between actions are so poorly designed that the sequences of actions prescribed by such relations are made unnecessarily difficult, or even impossible, to follow (as when dependency relations prescribe impossible tasks). Plans might also be *externally* ineffective when their purposes and sub-purposes are made unnecessarily difficult, or even impossible, to achieve. With this in mind, I believe it would be a mistake to exclude the possibility that even internally and/or externally ineffective plans might coordinate our activities. After all, as the mishaps in bureaucratic organizations suggest, we often act together by doing our parts in ill-conceived plans. There is no need, therefore, to impose any condition on plans (besides, of course, the tight explanatory link between plans and our joint conformity to them already included in my analysis) for such plans to provide for shared agency.³⁰

‘Shared’ in ‘shared agency’ might be thought to imply that every participant’s agency is involved to the same extent in shared activities. One might therefore worry that my analysis is unduly broad insofar as it considers our activities to be shared even when, for instance, I forcefully harness your agency through my threats and lead you to take part with me in the enactment of my plan. It has not been my intention, however, to suggest that shared activities never involve agency-undermining behaviors such as coercion or exploitation. Using Bratman’s example (2014, p. 102), British prisoners of war can be said to act together with Japanese troops when building the bridge over the River Kwai: though they would surely have acted differently if it were not for the background threats and coercion, their agential powers are involved in the enactment of the plan meant to build the bridge.³¹ In short, that my analysis leaves open the possibility that coercion or exploitation be involved in plan-coordination is no argument against the sufficiency of plan-coordination for shared agency. Quite the opposite, my analysis allows, as it should, for the often conflictual nature of our shared endeavors.

Here is a somewhat related worry. My plan-coordination analysis of shared agency is specifically designed to capture the fact that, as argued earlier (§4), shared intentions are not necessary for shared agency. Accordingly, it must be the case that shared intentions, on my analysis, are not necessary for the enactment of common plans.³² At this juncture, one might worry that my analysis is, again, unduly broad: if shared intentions are not necessary for plan-coordination, then it is possible, on my analysis, that a participant in a shared activity might engage in the form of strategic conduct that lies at the heart of traditional game theory and that shared agency theorists widely

²⁹ The following is inspired by Peter Graham (2011).

³⁰ To be sure, some plans are so radically defective that it is simply impossible to act in conformity to them. Given the first condition of my analysis of shared agency (namely, that our activities conform to some plan of action), it follows that such radically defective plans cannot provide for shared agency. This observation, however, is no argument against the sufficiency of my analysis. Indeed, that plans be not radically defective is not a further condition that should be added to the proposed analysis of shared agency. It is, rather, an entailment of my analysis: that our activities conform to a plan of action entails that conformity to that plan is possible and thus, that that plan is not radically defective.

³¹ There is an upper bound to the amount of agency-undermining behaviors compatible with shared agency: if I tie you up, push you into the trunk of my car, and then drive to NYC, your agency is not at all involved in our going to NYC, and therefore we do not count as going to NYC together (Bratman, 1992). Notice that cases of that flavor where the doing of your bit is merely the outcome of brute force are excluded by my analysis. On my view, indeed, that which might be shared are *activities*, and activities are doings to the production of which one’s agency is essential. As the enactment of my foregoing plan to go to NYC does not involve your agency in any way, nothing in that plan counts as an activity of yours. According to my analysis, thus, our going to NYC does not instantiate shared agency.

³² In the following section (§7), I show that that is indeed the case by describing agency-sharing mechanisms that dispense with shared intentions.

believe is in tension with acting together (e.g., Kutz, 2000, p. 77; Gold & Sugden, 2007, p. 111–117; Bratman, 2014, p. 5–6). That is to say: a participant's conduct, on my analysis, might be motivated by their expectations about what other participants might or might not do, as well as by individual intentions that best respond to such expectations.³³ For illustration, recall the narcissistic dancers who'd rather not share the audience's attention with the other dancers of the company and who accordingly all perform together while not sharing the intention to do so (§4). In this scenario, it is natural to suppose that each dancer's conduct is strategic, i.e., that each dancer's conformity to the choreography owes much to their attempt to best respond to the other dancers' conduct. After all, each dancer's conduct might indeed be guided by the thought that, no matter what the other dancers do, following the choreography rather than not maximizes the attention they are likely to receive from the audience. Yet, despite the dancers' strategic, best-response behaviors, my analysis classifies their following of the choreography as an instance of shared agency, and some might worry that that is a mistake.

In response to this worry, let me first observe that on my plan-coordination analysis, just as much as on popular accounts of shared agency, shared activities do not boil down to mere aggregations of strategic behaviors. I grant that on my analysis (unlike on most accounts), a participant's conduct in shared agency, considered on its own, might amount to strategic, best-response conduct. However, my analysis does not classify a collection of strategic behaviors as a shared activity unless these strategic behaviors are coordinated by a plan. The strategic dancers from the foregoing example dance together to the extent that each of their strategic dancing is coordinated by the same choreography. Second, it is simply false that strategic, best-response behavior is incompatible with shared agency. Approaching a shared activity in a strategic spirit is perhaps uncooperative—or at the least, not fully cooperative.³⁴ But clearly, it is a mistake to require of an agent that their conduct be fully cooperative for them to count as participating in shared agency. Those who coerce or exploit coparticipants behave uncooperatively; yet, as we have just seen, neither coercion nor exploitation is incompatible with shared agency. Similarly, then, strategic conduct, although less than fully cooperative, is compatible with shared agency. After all, our often conflictual social lives are rife with strategic participation in shared agency—a fact that my analysis captures.

Here is yet another worry. Margaret Gilbert (2013) has argued that our activities are shared only when we are mutually obligated to do our bit. Roughly, she believes that we do not genuinely act together unless I owe to you the performance of my part and you therefore have a special standing to demand compliance and rebuke me should I let you down. On Gilbert's view, then, plan-coordination falls short of shared agency.

But that is a mistake. The dancers of a decidedly experimental company might all deeply value their own artistic freedom, as well as that of their partners, and therefore agree to frame their collaboration as entailing neither the obligation to remain involved, nor the obligation to follow the choreography that emerged spontaneously during rehearsals. Yet, they will still succeed in giving their show together if each arrives at the right time and place for the performance and sticks to the non-obligatory dance moves of the choreography. This shows that what accounts for

³³ To be clear, the participants' conduct would also have to be plan-coordinated if they are to satisfy my analysis.

³⁴ Strategic behavior, in the game-theoretic sense that is my focus here, need not mean opportunistic (let alone, selfish) behavior. Behaving strategically, rather, means forming expectations about other agents' moves (which, therefore, are regarded as a given), and enacting individual intentions that best respond to such expectations. This form of strategic, best-response behavior, even if neither opportunistic nor selfish, might still fall short of cooperativeness: we might think, indeed, that cooperative agents do not simply best respond to their partners' behaviors but, instead, engage in a more robust manner with their partners' agency.

the sharedness of shared activities is not the obligation to play one's role (Bratman, 2014; Ludwig, 2016; Salomone-Sehr, 2023), but the fact that all participants' activities are coordinated by the same plan of action.³⁵

Besides plan-coordination, might it not be necessary that those involved in the enactment of a plan know—or at least believe—that they are so involved in order to count as acting together? In the most central cases of shared agency, perhaps. That said, we might have to allow for the possibility of shared agency under ignorance. To see this, suppose you are part of our dance company, and that you are such a narcissist that you will only do your best when alone on stage. The company's choreographer is aware of this. They also know that no one compares to you when you put your heart and soul into your practice. However, the choreography they envision requires more than just your performance. To achieve their artistic goals, the choreographer therefore resolves to lie to you and arrange the stage so that you remain unaware of the presence of the other dancers. On the day of the premiere, you are dancing together with the other dancers of the company even though you share the stage with them unknowingly: the act you are involved in is not a solo, but a tableau, whether you know it or not. This suggests that, perhaps, we should refrain from including any doxastic or epistemic condition in the analysis of shared agency.³⁶

Some might judge that in casting my net so wide, I have undermined the distinctiveness of shared agency. But the converse is true: the false belief that shared intentions are necessary for shared activities has distorted our appreciation of the proper explanandum of shared agency theory and led us to dismiss a range of cases as falling outside the scope of inquiry. My analysis, by contrast, handles the whole spectrum of activities done together, and not just the upper end of that spectrum where shared activities result from tightly knitted shared intentions and where participants satisfy strong epistemic conditions such as common knowledge. In particular, my analysis adequately accommodates instances where φ is done together in a robust sense, even though there is no description of φ that figures either in the content of some shared intention or in the content of some shared doxastic state.³⁷

³⁵ Gilbert should concede this point. After all, she agrees that agents might agree to proceed as if their shared endeavor involved no obligation (2013, p. 112). When that is so, what explains the sharedness of such endeavors, if not the common plan that coordinates the participants' activities?

³⁶ The foregoing coheres with my view that shared intentions are not necessary for shared agency. Earlier (§4), I showed that my intentions need not reflect the collective nature of the activity in which I participate. The point that emerges now is a generalization of this observation: on my view, that I act together with others need not be represented by my mental states, be they intentions, beliefs, or knowledge states. My view, for this reason, has obvious affinities with *externalist* accounts of social phenomena (Ritchie, 2020). I should insist that, even if I am inclined to count the foregoing dance scenario as a case of shared agency, I am willing to concede that such cases of shared activity under ignorance belong to the margins of shared agency.

³⁷ This shows (as helpfully observed by an anonymous referee) how much of a mistake it would be to treat 'shared agency' as a synonym of 'shared intentional activity' and the like (e.g., 'collective intentional action', 'joint intentional action', etc.). Notice that even when there is no description of φ that figures either in the content of some shared intention or in the content of some shared doxastic state, still is it the case, on my view, that for each participant, there is a description under which what they do is intentional (where these descriptions do not have overlapping contents). I am, after all, giving an analysis of shared *activities*, and activities are behaviors that are intentional under some description. In that respect, I am offering an analysis of shared agency rather than of collective *behavior*. Moreover, insofar as non-human animals do not seem to have the capacity for intentional behavior (of the sort that I am interested in anyway), the kind of social behavior in which non-human social animals (such as honeybees) engage in lies outside the scope of this article. With that being said, it seems to me that the conceptual apparatus that I invoke in my analysis of shared agency can certainly be tweaked to analyze collective behavior. More specifically, it might be thought, using this conceptual apparatus, that a collection of behaviors constitute a collective behavior if (and, perhaps, only if) these behaviors are coordinated by a single plan

That the sense of ‘together’ captured by my analysis is robust should be obvious by now. My analysis, indeed, straightforwardly satisfies the desiderata I argued for earlier (§3). Obviously, the foregoing analysis does not conflate multi-agent causation and shared agency: it is not just because individuals have caused some outcome together that they count as having enacted a common plan. Similarly, this analysis explains effortlessly what a common practical orientation amounts to: it is, simply, a *common plan of action*. Finally, this analysis captures the fact that constituent individual activities of a shared activity follow a common practical orientation non-coincidentally. As per the second condition of my plan-coordination analysis, things done together do not just conform to a common plan of action; rather, they are made to jointly conform to that plan by that plan. This excludes cases where conformity to a plan is merely coincidental.

Even so, you might perhaps still balk at the thought that what my analysis homes in on is a distinctive phenomenon of shared agency. After all, ‘shared agency’ is a bit of a term of art, one that some shared agency theorists might insist that we reserve for endeavors that enact shared intentions. At this point, however, the debate is merely terminological. Whichever name you might want to use, the phenomenon captured by my analysis is a pervasive phenomenon of our social life: one of acting together, in a robust sense, by virtue of plan-coordination.

Before moving on, let us pause to notice the important similarities between my plan-coordination analysis of shared agency and the collective goal account developed by Butterfill and Sinigaglia (Butterfill, 2016; Butterfill & Sinigaglia, 2023). On the collective goal account (2023), we act together if our actions have a collective goal, where that goal is an outcome and collective directedness to that outcome is cashed out in terms of there being a state, structure or situation which contributes to coordinating our actions in a way that is likely to result in that outcome’s occurrence. One first reason why my plan-coordination analysis and the collective goal account are similar is that there is a deep connection between the notions of plan and goal. As argued earlier, a plan involves a goal(s)—the goal(s) that the completion of the different steps of the plan is meant to achieve. Conversely, a goal might be considered to amount to a very underspecified plan.³⁸ Here is another important similarity. As we saw, one condition of Butterfill and Sinigaglia’s account is that there is a state, structure or situation that plays a role in coordinating the agents’ actions towards the achievement of some goal. Now, this condition resembles the second condition of my analysis, namely, that the plan contributes to explaining our joint conformity to it.³⁹ In light of this, it seems to be the case that whenever a collection of actions have a collective goal, where collective goal-directedness is understood along Butterfill and Sinigaglia’s account, that collection of actions also conform to a plan which figures in an explanation of these actions’ joint conformity to that plan—and vice versa. There must therefore be significant extensional overlap between my plan-coordination analysis of shared agency and Butterfill and Sinigaglia’s collective goal account, and so the latter should be recognized as a very important precursor of the former.

(where the plan in question is not a plan of *action* but rather what we might call a *behavioral plan*—a plan that specifies a blueprint for a collection of behaviors). I thank an anonymous referee for suggesting that I flag how the tools developed in this article might fruitfully illuminate the sort of collective behaviors in which some non-human animals engage in.

³⁸ Here again, I am exploiting the idea that plans are susceptible of varying degrees of specificity. See endnote 24 above.

³⁹ Butterfill’s 2016 discussion of collective goals (2016) differs in some respects from the collective goal account presented by Butterfill and Sinigaglia (2023). One difference is that, unlike Butterfill and Sinigaglia (2023, p. 9), Butterfill’s 2016 article does not cash out directedness to a goal in terms of coordination; rather, it invokes coordination as a possible way in which goal-directedness might have a collective nature (2016, p. 316)—which suggests that, in Butterfill’s 2016 article, coordination neither characterizes goal-directedness, nor is thought to be necessary for collective goal-directedness. Another important difference is that it is not entirely clear how committed Butterfill ultimately is, in his 2016 article, to the truth of the claim that collective goals suffice for shared agency (2016, p. 367).

There are nevertheless some differences between my plan-coordination analysis and Butterfill and Sinigaglia's account. The extensional overlap between my analysis and their account might not be perfect. In my analysis, I insist that the explanatory link between the plan and our joint conformity to it must be *tight*. By contrast, Butterfill and Sinigaglia merely argue that the state, structure or situations they mention in their account must play *a* role in coordinating the agents' actions towards the achievement of some goal. Because they do not require that the role in question is to *tightly* coordinate the agents' actions, their account might be vulnerable to deviant causal chains⁴⁰—a problem that my analysis is designed to avoid. (I discuss deviant causal chains in greater detail in §7.1.)

Second, Butterfill and Sinigaglia, as seen previously, are open to the idea that the conditions of the collective goal account might not be necessary for shared agency. It is my view, by contrast, that the plan-coordination analysis provides necessary conditions for shared agency. Even if there were perfect extensional overlap between their account and my analysis, Butterfill and Sinigaglia would still be open to there being extensional daylight between my analysis and a full characterization of shared agency.

Third, unlike their account (and for that matter, unlike Salomone-Sehr's [2022]),⁴¹ my analysis does not rely on the somewhat vague notion of coordination. Coordination by a plan, in my plan-coordination analysis, is cashed out in terms of their being a plan to which agents conform to, and that contributes to explaining their joint conformity to it.

Lastly, it seems to me that my plan-coordination analysis draws attention to the most salient feature of things done together, namely, the enactment of a common plan—rather than the pursuit of a collective goal. Consider greeting rituals: it is certainly far easier to notice and describe the piece of social choreography that they involve than the goal(s) they serve, as Ruth Millikan herself has observed (2004, p. 5). Besides, there are instances of shared agency where the language of goal, while remaining adequate, nevertheless seems idle. Consider for instance children playing a hand-clapping game: What exactly is the collective goal of the game, if not to conform to the steps of a plan—the plan provided by the hand-clapping choreography?^{42,43}

With this, I conclude my plan-coordination analysis of the concept of shared agency. Let us now turn to mechanism explanation, the second half of my account.

7 | TWO GRADES OF MINIMALISM

In this section, I describe some of the mechanisms that enable the formation and enactment of common plans, and hence that provide for shared agency. Sometimes, plan-coordination goes by

⁴⁰ It should be noted that Butterfill and Sinigaglia perhaps do not intend the state, structure or situation involved in their collective goal account to play a *causal* role in coordinating the agents' actions towards the achievement of some goal. In other words, Butterfill and Sinigaglia's collective goal account might perhaps not be a causal account of shared activity—in which case the collective goal account would avoid the problem of deviant causal chains (as an anonymous referee helpfully observed).

⁴¹ Here is how the present article relates to my 2022. In contrast with my 2022, the present article does not restrict attention to *cooperative* shared activity; rather, it aims to account for shared activity of any sort, whether cooperative or not. The present article also develops further the analysis of shared activity that my 2022 initially formulated in terms of coordination (see my discussion of the Coordination Condition [Salomone-Sehr, 2022, p. 434-438]).

⁴² I said above that plans involve goals. What this example reveals is that sometimes, the goal of a plan is nothing more than enacting the steps of the plan.

⁴³ I thank an anonymous referee for inviting me to explain how my analysis relates to Butterfill and Sinigaglia's.

way of shared intentions: when that is the case, our plan figures in the content of our shared intentions, and plan-coordination, accordingly, is just a matter of enacting this jointly intended plan.

Besides shared intentions, however, there are at least two additional families of agency-sharing mechanisms. In this section, I sketch an account of both. In so doing, my aim is to see how far we can go in making plan-formation and plan-coordination *not* depend on agents' psychologies. That is, I explore how *minimalist* the mechanisms that provide for acting together can be.

In passing, we should note that this section can be read as providing further support in favor of my plan-coordination analysis. My discussion here will indeed make good on my earlier claim that my plan-coordination analysis provides a generic characterization that can be applied to the whole spectrum of shared activities—including the minimalist end of that spectrum.⁴⁴

7.1 | Moderate minimalism: planner-based shared activities

The first kind of mechanism features a central planner (e.g., a choreographer). The planner forms a plan through an episode of practical reasoning, and makes sure to decompose that plan into subplans. Agents of some collection are then each assigned a subplan—that is, they are assigned a *role* in that plan. They then enact the plan by following their respective roles in that plan.

These agents can conform to the planner's plan by sharing the intention to enact it, and then by acting on that shared intention: when following each of their roles, therefore, they do so with a view to achieving the plan of which these roles form a part. But I believe there are at least two alternative, less psychologically demanding ways through which agents might conform to their plan. Each agent might *merely* intend to play their respective role in the plan, as when I commit myself to doing my bit on the assembly line but do not care whether what I do contributes or not to the production plan's success—let alone whether the goals of the production plan will be successfully achieved. Alternatively, the collection of agents might be nudged into completing the tasks involved in their roles. In some subway stations, footprint stickers stuck on the ground indicate paths which, when followed, ensure a smooth traffic of passengers. Most of us, I take it, follow such stickers mindlessly, that is, without forming the intention to walk in the direction indicated by the stickers: rather, our intention might just be to cover the footprint stickers with the sole of our shoes. Yet, acting on that intention suffices to enact a plan meant to avoid collisions and congestion. Note that cases where behavioral nudges ensure a collection of activities' joint conformity to some plan offer an illustration of shared activities where agents might share their agency unknowingly: commuters following footprint stickers need not realize that in so doing, they are participating in a common plan.

As argued previously (§6), shared agency takes more than conforming to a common plan. It requires, additionally, that the plan explains our joint conformity to it. How might a planner's plan figure in an explanation of the enactors' joint conformity to it?

First, the plan⁴⁵ must be causally efficacious in moving the enactors to act as required by the plan. But that is not enough. Suppose I am a teenager and you are my mom. You are fed up with doing all the chores, and so you form the plan that my sister and I do our part by cleaning our

⁴⁴ I thank an anonymous referee for making this observation.

⁴⁵ More precisely, that which must be causally efficacious is not so much the plan (which, qua abstract object, might well be causally inefficacious, see endnote 22 above) as the mental representations of the planner in which the plan figures.

respective rooms. You communicate this plan to dad, just to warn him. Dad is an old-fashioned man who pays little attention to fairness in domestic labor, and therefore does not engage much in your plan. Hearing you talk about cleaning rooms, however, gives him an idea: he could force me to clean my room as punishment for the mess I made the other day when I used his laptop without his authorization. He could, and he does: before you could tell my sister and me about your plan to divvy up domestic chores differently, he orders that I clean my room. As a result, here I am, cleaning my room. And since my little sister does everything that I do, she is now cleaning hers too.

In this admittedly farfetched scenario, there is a causal chain that starts with your plan and that goes all the way to my sister and I acting in joint conformity to your plan.⁴⁶ However, it seems wrong to say that we are enacting your plan, hence engaged in doing domestic chores together. Instead, my behavior enacts dad's plan, and my sister's behavior enacts her own plan to copy whatever it is that I do.

To solve this version of the problem of deviant causal chains,^{47,48} I suggest adding these two conditions to the foregoing account. First, the planner's plan must be *self-referential*:⁴⁹ the plan does not just specify how a collection of agents is supposed to act; additionally, it is part of the content of the plan that the plan⁵⁰ itself be causally involved in the production of the enactors' joint conformity to it. Second, the plan must be involved in the production of the enactors' joint conformity to it not just in any possible way, but in the way that the plan stipulates.

These additional conditions keep at bay deviant causal chains and ensure that the explanatory link between the plan and our joint conformity to it is tight, as required by my plan-coordination analysis.⁵¹ Going back to the previous example, your plan is not merely that my sister and I clean our room. Additionally, it is part of your plan that your very plan be involved in a certain way in bringing it about that my sister and I clean our respective rooms. And whatever this way is, it is certainly not your intention that dad, as he is listening with only half an ear to the new division of domestic labor you have all sorted out, hits upon the idea to use the cleaning of my room as punishment for my past mischief—neither is it your intention that my sister winds up cleaning her room too because she does everything I do.

With these additional conditions, is my account now too demanding? I do not think so. For one thing, both conditions reflect good practices that planners adopt when in the business of coordinating others' actions. Planners do not just specify how collections of agents are supposed to act. They also deliberate about how to make their plans effective; thus, they settle on a plan

⁴⁶ *Joint* conformity is here ensured by my sister's tendency to copy everything I do.

⁴⁷ The problem here is one of tightness of the explanatory relation between the plan and our joint conformity to it.

⁴⁸ Like Shapiro, I argue that shared intentions are not necessary for shared agency (Shapiro, 2014). And like him still, I offer an account of shared agency on which acting together requires that our activities conform to a plan and that plan explains our conformity to it (2014, p. 277). There are, however, some notable differences between his account and mine (some of which I flagged earlier in endnote 21). One such notable difference is that Shapiro has not observed that the explanation that links the plan to our conformity to it must be tight. His account therefore fails to address deviant causal chains and, as a result, is laid open to the charge of overgeneration. For instance, Shapiro's account, as stated, would seem to mistakenly classify the foregoing domestic chores scenario as an example of shared agency. I thank an anonymous referee for asking me to clarify how my account compares to Shapiro's.

⁴⁹ See Gilbert Harman (1976, p. 444) and Ludwig (2016, p. 88-97).

⁵⁰ More precisely, a representation of the plan.

⁵¹ See §6.

that describes with some specificity how that very plan is supposed to lead to its own achievement. Moreover, the thought that my account of the foregoing agency-sharing mechanism is too stringent misses the point of the task I set out to complete here. The point of this section is to describe mechanisms for shared agency, i.e., conditions which, when satisfied, suffice for plan-formation and plan-coordination, and hence for acting together. Accordingly, the focus here is not on necessary conditions, but sufficient ones.

Notice that in the agency-sharing mechanism I described, shared intentions are not necessary. Insofar as it does not require the rich web of interpersonal psychological structures that the sharing of intentions necessitates, this mechanism is, thus, minimalist. Nevertheless, it still crucially involves a planner and their planning activity: in this mechanism, the content of the plan is determined by the planner's practical reasoning, and whether a collection of activities are coordinated by the same plan depends (partly) on the content of the planner's plan. For these reasons, this mechanism is *moderately* minimalist.

7.2 | Strong minimalism: selective history-based shared activities

Suppose we now try to remove the planner. Can plans be formed and can they coordinate our activities even in the planner's absence? Put differently, might there be strongly minimalist mechanisms that enable shared agency without the sharing of intention *and* without the multi-agent enactment of a planner's plan?

The idea of plans without planners, perhaps odd at first, is in fact a familiar one when interpreted with the tools of evolutionary theory (Graham, 2011). In this picture, the instantiation of a plan by some item or other (e.g., an organism) is the outcome of a history of selection: the plan in question specifies how a selected item has been operating for it to be selected. One major insight of evolutionary theory is that selection does not necessitate anyone's intentions. In fact, smart plans might arise through the combination of rather unintelligent processes.

Here are two applications of this package of ideas to shared agency theory.

7.2.1 | Trial and Error⁵²

Imagine a bunch of passengers getting off a densely packed NYC subway car at Union Square. These passengers spontaneously form a line that inches towards the subway doors. Given how crowded that car is, there is almost no space between two consecutive passengers in that line. The risk of collision is therefore pretty high: if one passenger moves forward too fast, they will bump into the passenger ahead of them; if they are too slow, the passenger behind them will step on the back of their shoes. All these passengers, accordingly, are badly in need of a common plan for the coordination of the speed at which they each move towards the exit.

My suggestion is that the content of that common plan might be determined by individual trial and error. A passenger standing in that spontaneously formed line of exiting passengers might at first move forward too fast, then realize they should immediately slow down to avoid bumping into the passenger ahead of them, until it becomes clear that they are now going too slow as they can feel the passenger behind them pushing forward. As exiting passengers each engage in

⁵²This subsection has greatly benefitted from an anonymous referee's insightful comments.

similar trial and error,⁵³ it seems possible that, after several failed attempts, all such passengers end up aligning the speed at which they creep towards the exit, thus enabling them to leave the train while avoiding major collisions. When that happens, their repeated attempts can be said to have homed in on a common plan of action. This plan, in turn, explains the passengers' sustained conformity to it: once they start conforming to that plan by moving forward at the speed that that plan prescribes, it becomes costly to deviate from that plan, for they would then run the risk of stepping on each other's feet. The plan, therefore, explains their sustained conformity to it.⁵⁴

It might be objected that the coordination that the commuters have achieved when exiting the subway car falls short of *shared* agency. Some readers might think, indeed, that the subway scenario provides an illustration of individual agency in a social context. On this line of thinking, the scenario in question does not involve a common plan, but rather the passengers' individual plans to exit the subway car while avoiding collisions.⁵⁵

In response, I grant that such individual plans do play an important role, especially at the initial stages of the commuters' attempts to coordinate. However, they cannot be the full story. A full explanation of the commuters' conduct must advert to the common plan that their initially failed attempts, lo and behold, have ended up selecting. After all, there is a whole range of different speeds at which the line of passengers could have moved forward towards the subway doors while avoiding collisions. It so happens, however, that the passengers, through trial and error, have homed in on *some* specific speed. An adequate explanation of the passengers' sustained conformity to that speed *must* cite that very speed and, in particular, the fact that deviating from that speed would be likely to result in collisions. Indeed, only citing the passengers' individual plans to get off without collision will not provide an adequate explanation of the passengers' sustained conformity to that very speed, given that such individual plans are at first compatible not just with that speed, but with a range of speeds. From that discussion, it follows that in this scenario, avoiding a collision is something the passengers do *together* in the robust sense of 'together' captured by my plan-coordination analysis of shared agency: they follow a plan, and that plan figures in an explanation of their joint conformity to that plan.⁵⁶ True, the common plan that coordinates the commuters' moves is rather unsophisticated insofar as it only prescribes the speed at which the commuters must inch towards the exit. But it is easy to see how a variation of this example could involve a more complex plan. For instance, the subway car has perhaps multiple doors and trial and error might help the commuters home in on not just a uniform speed, but also an assignment of commuters to one of the many doors.

Perhaps you might still think, at this point, that the foregoing subway example falls short of shared agency because, for any of these exiting passengers, it might not matter much at all whether the other passengers had rather stayed in the subway car, tripped, or got blocked.⁵⁷ We have seen

⁵³ It is natural to suppose that in this example, each commuter's trial and error is an instance of strategic, best-response behavior—that is, the kind of behavior that shared agency theorists believe is in tension with shared agency. As we saw before (§6), strategic behavior and shared agency are in fact compatible.

⁵⁴ Above (endnote 22), I said that one might wonder how a plan (that is, an abstract object) can figure in an explanation of our conformity to it. I believe there is nothing mysterious in the explanation described in the main text.

⁵⁵ I thank an anonymous referee for pressing this objection.

⁵⁶ I observed earlier that, on a natural description of the case, each commuter engages in strategic, best-response behavior (see endnote 53). To the extent that their moves end up coordinated by a plan (i.e., the plan that emerged through trial and error), what the commuters do does not boil down to a mere aggregation of strategic behaviors.

⁵⁷ I am here borrowing an anonymous referee's helpful language.

already, however, that the Shared Intention Thesis is false (§4): participants in a shared activity need not be committed to the collective nature of that activity. It is false, in turn, that for the subway example to involve shared agency, every passenger must intend the enactment of a course of action where, in addition to getting off the subway car, other passengers, too, get off that subway car.

Before moving on, notice that trial and error is structurally similar to Darwinian selection (Godfrey-Smith, 2012). First, each trial introduces variation. Second, successful attempts form the basis for subsequent cycles of trial and error. Third, trial and error is a fairly rudimentary process, yet likely to home in on smart solutions. These remarks together with the subway example therefore suggest that a roughly Darwinian process might account for the plannerless formation of plans, as well as the plannerless plan-coordination of our activities.

7.2.2 | Selection of Plan-Types

Let us now zoom out and explore how more or less Darwinian processes might lead to the formation of plan-types that might then be deployed not just in one single interaction, as in the previous example, but across different interactions. To fix ideas, consider greeting rituals (Millikan, 2004, p. 20). (Millikan suggests other plausible examples such as behaviors engaged in during shared meals, styles of interactions in the market, courting behaviors, and manners in which people accept or politely refuse offers.) Greeting rituals obviously involve a plan, that is, a specification of the range of expressions that can be said and the sort of gestures that are usually expected. True, this kind of social choreography can be the result of intentional design (which was the case when greetings rituals were recently intentionally reshaped to reduce the risk of infection by SARS-CoV-2). But it seems possible that greeting rituals might emerge, spread, and coordinate our interactions even in the absence of a planner.

Under which conditions would this possibility materialize? More generally speaking, under which conditions might plan-types emerge, spread, and coordinate our activities by way of roughly Darwinian mechanisms?

Here is one such set of conditions (among perhaps other such sets). First, there might be a population of items where there is variation in character and where the selected-for features of these items determine the content of not just any sort of plan, but of a common plan *of action*. What sort of items fits that description? One plausible answer, I believe, is *interactions*, i.e., the face-to-face encounters in which much of our rich social lives consists.⁵⁸

Second, there might be some way to make sense of the idea that some styles of interactions are more *fit* than others—for instance, because they are more enjoyable, more readily remembered, or easier to execute, and hence more likely to spread.

Third, there might be some mechanism through which styles of interactions spread differentially better depending on their relative fitness. In biology, *reproduction* plays such a role and it might seem that interactions are not the sort of things that can beget other interactions. There are, at least, two responses to this challenge. One is to tighten the analogy between interactions and the sorts of populations of individuals studied by biologists. Take for instance couples dancing

⁵⁸ Cultural evolution theorists use the term ‘cultural variant’ (Godfrey-Smith, 2012; Richerson & Boyd, 2008) or ‘memes’ (Dawkins, 2016) to describe the hotchpotch of cultural items that might undergo processes of cultural selection. This hotchpotch includes behaviors, skills, values, attitudes (Richerson & Boyd, 2008, p. 63), and artifacts (Sterelny, 2006). Cultural evolution theorists seem to have neglected the possibility that interactions might count as cultural variants too.

the tango. Perhaps the dancing style of every couple is the faithful imitation of some other couple's own style. If that were the case, tango interactions would form lineages of the sort observed among organisms. Another response is to relax the idea that each interaction have a small number of parent interactions. For instance, we might tend, to an extent, to engage in the kind of greeting ritual that strikes us as the most prevalent among our social circles. In other words, a bias for conformity might affect, to an extent, the spread of greeting rituals. As a result of this bias, each greeting would count as having as many sources as greeting rituals previously observed, thus undermining the idea that such sources might be analogous to parents.⁵⁹ Given that Darwinian evolution paradigmatically involves reproduction by a limited number of parents (Godfrey-Smith, 2009), relaxing the idea that interactions have a small number of parents would amount to diluting the Darwinian character of the evolution that might affect interactions. This sort of evolution might still remain Darwinian (for instance if the bias for conformity only has a limited influence).

If these conditions were satisfied, interactions would lend themselves to the kind of search processes that organisms undergo, and the features of these interactions that have explained these interactions' selection would, thus, determine the content of what would be tantamount to plans of action.

So far, I have explained how roughly Darwinian processes might lead to the formation of plan-types.⁶⁰ Now, how might such plan-types explain the fact of our joint conformity to them? Here is one mechanism: as in the case of trial and error developed above, such plan-types, once selected, might create incentives favoring conformity to them. If deviation from a plan-type is costly, then that plan-type explains our enduring joint conformity to it.

To illustrate, let us speculate and suppose that greeting rituals did arise through a process of the kind I just described. Imagine now that you and I greet each other at the office on a Monday morning. We know that social choreography and because deviance from it is disapproved of, we both conform to it as we bump into one another in the elevator. In that example, you and I engage in shared agency even though neither the determination of our plan, nor the plan-coordination of our activities, have required either shared intentions or a planner.^{61, 62}

⁵⁹ Introducing a bias for conformity would also undermine the idea that styles of interaction must be fit to be selected.

⁶⁰ Note that the plannerless emergence of plans need be neither wholly Darwinian (as seen earlier when I discussed the bias for conformity), nor perhaps Darwinian *at all* (say, if the evolution of some form of interaction is overwhelmingly driven by a bias for conformity, and hence independent of fit). In such cases where evolution is overwhelmingly driven by a bias for conformity, the content of the evolved plan of action might perhaps be thought of as the features of some form of interaction to which conformist conduct latches on to and tends to mimic. Provided that the plan in question explains joint conformity to it (for instance if there are incentives against deviance), it would seem to provide for plan-coordination and, thus, for shared agency. In the foregoing, I have focused on Darwinian evolution only because it is a familiar mechanism of plannerless plan-formation. Once that mechanism is on the table, it is easy to see how that mechanism can be tweaked and remain plannerless (while perhaps losing its Darwinian character).

⁶¹ It should be stressed that the fact that a greeting ritual has evolved over time without intelligent design does not mean that we might not have a shared intention to greet each other: we could indeed share the intention that we both engage in that specific greeting ritual. Interestingly, this shows that some of the plans that might figure in the content of shared intentions need not be plans of anyone's design.

⁶² To motivate the need for a minimalist account of shared agency, I earlier gave the example of a housewife and husband enacting interlocking gender roles together (§2). Speculating a bit, we might think that such roles might have emerged through some evolutionary mechanism or other in part in virtue of the coordinative function that they serve (despite their often-oppressive nature). For a formal elaboration of this idea, see O'Connor (2019, especially chapter 4).

8 | CONCLUSION

I have argued that we need a minimalist conception of shared agency, one that does not deem, mistakenly, that shared intentions are necessary for shared agency. I have also argued for three adequacy criteria that any account of shared agency, however minimalist, should satisfy. The minimalism and these criteria have created a puzzle that I have subsequently solved in two stages: first, I have analyzed shared agency in terms of the notion of plan; second, I have described two agency-sharing mechanisms that do not rely on shared intentions. One major virtue of my account of shared agency is that it captures the whole spectrum of shared activities, from the cases where participants are closely knitted, and where the common plan figures accordingly in a shared intention of theirs, to the cases where the ties between participants are much more attenuated, and where, thus, the common plan might not figure in anyone's psychology.

ACKNOWLEDGMENTS

My deepest thanks to Miranda Fricker for her extremely generous support and thorough comments on successive drafts. I would also like to thank Françoise Arqué-Sehr, Peter Baumann, Olle Blomberg, Sacha Bourgeois-Gironde, Michael Bratman, Thomas Brouwer, Rodrigo Díaz, Kirk Ludwig, Charles Mills, Jennifer Morton, Elisabeth Pacherie, David Papineau, Marie-Françoise Péneaud, Matthew Rachar, Katherine Ritchie, Abraham Roth, Ren Schaeffer, Eric Shoemaker, Nicholas Taylor, Robert Williams, as well as the audiences at the 2020 Social Ontology Conference (Neuchâtel, online) and the 2022 Eastern Meeting of the American Philosophical Association (Baltimore, online). Lastly, I express my gratitude to an anonymous referee for *Philosophy and Phenomenological Research*. I believe their careful comments have considerably strengthened my arguments.

This article has received financial support from the Canada Research Chair on Epistemic Injustice and Agency.

ORCID

Jules Salomone-Sehr  <https://orcid.org/0009-0009-3588-6124>

REFERENCES

- Alonso, F. M. (2009). Shared Intention, Reliance, and Interpersonal Obligations. *Ethics*, 119, 444–475. <https://doi.org/10.1086/599984>
- Asarnow, S. (2020). Shared Agency Without Shared Intention. *The Philosophical Quarterly*, 70, 665–688. <https://doi.org/10.1093/pq/pqaa012>
- Brand, M. (1986). Intentional Actions and Plans. *Midwest Studies in Philosophy*, 10, 213–230. <https://doi.org/10.1111/j.1475-4975.1987.tb00541.x>
- Bratman, M. (1987). *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press.
- Bratman, M. (1992). Shared Cooperative Activity. *The Philosophical Review*, 101, 327–341. <https://doi.org/10.2307/2185537>
- Bratman, M. (2014). *Shared Agency: A Planning Theory of Acting Together*. Oxford: Oxford University Press.
- Butterfill, S. (2012). Joint Action and Development. *The Philosophical Quarterly*, 62, 23–47. <https://doi.org/10.1111/j.1467-9213.2011.00005.x>
- Butterfill, S. (2016). Joint Action: A Minimalist Approach. In J. Kiverstein (Ed.), *The Routledge Handbook of Philosophy of the Social Mind* (pp. 357–369). Milton Park: Routledge.
- Butterfill, S., & Sinigaglia, C. (2023). Towards a Mechanistically Neutral Account of Acting Jointly: The Notion of a Collective Goal. *Mind*, 132, 1–29. <https://doi.org/10.1093/mind/fzab096>

- Chant, S. R. (2006). The Special Composition Question in Action. *Pacific Philosophical Quarterly*, 87, 422–441. <https://doi.org/10.1111/j.1468-0114.2006.00270.x>
- Dawkins, R. (2016). *The Selfish Gene*. Oxford: Oxford University Press.
- Fricker, M. (2016). What's the Point of Blame? A Paradigm Based Explanation. *Nous*, 50, 165–183. <https://doi.org/10.1111/nous.12067>
- Fricker, M. (2019). Forgiveness—An Ordered Pluralism. *Australasian Philosophical Review*, 3, 241–260. <https://doi.org/10.1080/24740500.2020.1859230>
- Garcia, M. (2021). *We Are Not Born Submissive: How Patriarchy Shapes Women's Lives*. Princeton: Princeton University Press.
- Gilbert, M. (2013). *Joint Commitment: How We Make the Social World*. Oxford: Oxford University Press.
- Godfrey-Smith, P. (2009). *Darwinian Populations and Natural Selection*. Oxford, New York: Oxford University Press.
- Godfrey-Smith, P. (2012). Darwinism and Cultural Change. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 2160–2170. <https://doi.org/10.1098/rstb.2012.0118>
- Gold, N., & Sugden, R. (2007). Collective Intentions and Team Agency. *The Journal of Philosophy*, 104, 109–137. <https://doi.org/10.5840/jphil2007104328>
- Graham, P. J. (2011). Intelligent Design and Selective History: Two Sources of Purpose and Plan. In J. L. Kvanvig (Ed.), *Oxford Studies in Philosophy of Religion Volume 3* (pp. 67–88). Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199603213.003.0005>
- Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Harman, G. (1976). Practical Reasoning. *The Review of Metaphysics*, 29, 431–463.
- Khader, S. J. (2019). *Decolonizing Universalism: A Transnational Feminist Ethic*. Oxford: Oxford University Press.
- Knowles, C. (2019). Beauvoir on Women's Complicity in Their Own Unfreedom. *Hypatia*, 34, 242–265. <https://doi.org/10.1111/hypa.12469>
- Kutz, C. (2000). *Complicity: Ethics and Law for a Collective Age*. Cambridge: Cambridge University Press.
- List, C., & Pettit, P. (2011). *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.
- Ludwig, K. (2016). *From Individual to Plural Agency: Collective Action: Volume 1*. Oxford: Oxford University Press.
- Millikan, R. G. (2004). *Varieties of Meaning*. Cambridge, MA: MIT Press.
- O'Connor, C. (2019). *The Origins of Unfairness: Social Categories and Cultural Evolution*. Oxford: Oxford University Press.
- Pacherie, E. (2013). Intentional Joint Agency: Shared Intention Lite. *Synthese*, 190, 1817–1839. <https://doi.org/10.1007/s11229-013-0263-7>
- Paternotte, C. (2014). Minimal Cooperation. *Philosophy of the Social Sciences*, 44, 45–73. <https://doi.org/10.1177/0048393112457428>
- Pettit, P., & Schweikard, D. (2006). Joint Actions and Group Agents. *Philosophy of the Social Sciences*, 36, 18–39. <https://doi.org/10.1177/0048393105284169>
- Rachar, M. (2023). Conditional Intentions and Shared Agency. *Nous*. Advanced online publication. <https://doi.org/10.1111/nous.12452>
- Richerson, P. J., & Boyd, R. (2008). *Not By Genes Alone: How Culture Transformed Human Evolution*. Chicago: University of Chicago Press.
- Ritchie, K. (2020). Minimal Cooperation and Group Roles. In A. Fiebich (Ed.), *Minimal Cooperation and Shared Agency* (pp. 93–109). Berlin: Springer International Publishing. https://doi.org/10.1007/978-3-030-29783-1_6
- Roth, A. S. (2004). Shared Agency and Contralateral Commitments. *The Philosophical Review*, 113, 359–410. <https://doi.org/10.1215/00318108-113-3-359>
- Salomone-Sehr, J. (2022). Cooperation: With or without Shared Intentions. *Ethics*, 132, 414–444. <https://doi.org/10.1086/716877>
- Salomone-Sehr, J. (2023). Shared Agency and Mutual Obligations: A Pluralist Account. *The Philosophical Quarterly*, 73, 1120–1140. <https://doi.org/10.1093/pq/pqad041>
- Searle, J. (1990). Collective Intentions and Actions. In P. R. Cohen, J. Morgan, & M. Pollack (Eds.), *Intentions in Communication* (pp. 401–415). Cambridge, MA: MIT Press.
- Shapiro, S. J. (2014). Massively Shared Agency. In M. Vargas & G. Yaffe (Eds.), *Rational and Social Agency: The Philosophy of Michael Bratman* (pp. 257–293). Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199794515.003.0011>

- Shepherd, J. (2021). *The Shape of Agency: Control, Action, Skill, Knowledge*. Oxford: Oxford University Press.
- Sterelny, K. (2006). Memes Revisited. *The British Journal for the Philosophy of Science*, 57, 145–165. <https://doi.org/10.1093/bjps/axi157>
- Velleman, J. D. (2015). *The Possibility of Practical Reason*. Ann Arbor: Michigan Publishing.

How to cite this article: Salomone-Sehr, J. (2023). How to be minimalist about shared agency. *Philosophy and Phenomenological Research*, 1–24.
<https://doi.org/10.1111/phpr.13030>