Conditional Intentions and Shared Agency

Matthew Rachar

Abstract

Shared agency is a distinctive kind of sociality that involves interdependent planning, practical reasoning, and action between participants. Philosophical reflection suggests that agents engage in this form of sociality when a special structure of interrelated psychological attitudes exists between them, a set of attitudes that constitutes a collective intention. I defend a new way to understand collective intention as a combination of individual conditional intentions. Revising an initial statement of the conditional intention account in response to several challenges leads to a specification of the properties these intentions need to have in order to be genuine commitments. I then show how a structure of conditional intentions with these properties settles a collection of agents on engaging in social interactions that display all the features typically associated with shared agency.

When we act together, we depend on one another. If I aim at doing something together with you, I must see you as likewise aiming to do that thing with me. When each of us sees the situation this way and behaves accordingly, we do more than act in parallel, or do the best we can given what we think the other will do. We genuinely share agency. There is something we can properly speak of as our action. This paper offers a new way to understand the kind of interdependence involved in shared agency.

It is commonly held that the distinction above is explained by collective intention. And it has been suggested, although not widely accepted, that we can construct collective intentions by way

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. Noûs published by Wiley Periodicals LLC.
of combinations of individual conditional intentions towards an action of the same kind. The rough idea is that one way for us to form a collective intention to go for a walk, for example, is for one of us to say “I’ll go for a walk if you will”, and the other to say, “Well, then I will too”. Challenges to this idea have come from those who hold that the specific kinds of conditions on these intentions make them not really intentions at all, that, even if they are genuine intentions, they do not play the characteristic roles of collective intention, and that they can’t explain the interpersonal obligations associated with acting together.

The conditional intention proposal, on a particular construal, has the resources to respond to these charges. My aim here is to defend the coherence of this important kind of collective intention by developing that construal and response. First, I introduce the features of intention that make it a valuable component in constructing an explanation of this kind of social interaction, as well as a general strategy for identifying collective intention and an initial statement of the conditional intention account. Then, I address the first challenge by specifying what kind of conditions such intentions need to have in order to be genuine commitments. They need to be reciprocal, so that the conditional on each is the matching intention of the other participants. They need to be internal, in the sense that the other’s intention is part of the content of one’s own intention. And they need to be non-precautionary, which means that the intender does not see their partner’s intention as an unfortunate circumstance around which they must plan. I also consider whether they need to be elective, that is, something the agent chooses rather than accepts as necessary. After clarifying the nature and importance of the form of sociality at issue, I defuse the last two challenges by showing how combinations of these conditional intentions play the roles of collective intention, and how they may be combined with a background normative theory to explain the kinds of interpersonal obligations some think are involved in acting together.

The key to these explanations is that intentions with those three properties involve a special kind of interdependence. It is not just that each individual intention is conditional on the other intentions, but that each intention is conditional on the others’ intentions being conditional on it. This kind of interdependence is what settles each participant on taking their cue from the others, what settles them on genuinely sharing agency.

1 THE CONDITIONAL INTENTION ACCOUNT OF COLLECTIVE INTENTION

The central building block for constructing collective intention is participatory intention. A participatory intention is an individual intention directed towards our behaviour. There are some general features of intention that make it well-suited for this role. But these features also lead to a problem for the relationship between partners. The prospect of solving this problem is one motivation for appealing to conditional intentions.

Forming an intention is something that settles an agent’s deliberative question about what to do. There are two components to this kind of settling. The first is causal. Intentions tend to cause behaviour in their agents, and in order to be successful, they must actually cause the behaviour

---

1 This assumption is common in the literature (Bratman, 2014; Alonso, 2009; Tuomela, 2005; Ludwig, 2016). Roth (2014) calls it the “Intention Thesis” and provides a defense. It is denied by Gilbert (2013).

2 The following characterization is meant to pick out some of the main points from the debate, focusing on the common features of intention emphasized in Bratman (1987) and Velleman (2015).
they specify. The second is representational. An intention must represent itself to the agent as resolving the matter, at least for the time being.\textsuperscript{3}

Intention also prompts and constrains further deliberation in various ways, because of the norms and corresponding dispositions of practical reason. These rationality norms are means-end coherence, agglomerativity, stability, and consistency.\textsuperscript{4} Consistency requires that intentions be internally consistent and consistent with one’s beliefs. Agglomeration requires that intentions be combinable and meet the consistency norm. Means-end coherence requires that intentions in favour of ends engage a demand to choose means, as that becomes required. Stability puts pressure on retaining intentions over time, although it does allow for reconsideration and revision in certain circumstances. Because of these norms and our agential make-up, which involves dispositions to adhere to them, having an intention means being disposed and rationally required to resist temptations, filter out inconsistent plans, and find solutions to the means-end problems posed in execution. This allows intention to play its characteristic role in our lives of unifying our individual agency over time by settling deliberative questions about what to do.

The search for collective intention is the search for something that plays these roles with respect to the coordination and unification of behaviour, plans, and deliberation across people. The general methodology then is to find a suitable translation of the roles of intention in organizing individual agency over time to roles for collective intention in organizing agency across people and then to identify a structure that plays those translated roles. I focus on a psychological structure that features individual intentions and the associated rationality norms. I come back to the issue of a suitable translation in §3. For now, I note two limitations of this approach. I do not develop an argument to show that the structure I propose is the only structure that would realize these roles, nor do I attempt to show that it is necessary that such a structure include individual intention.\textsuperscript{5} Nonetheless, the nature of intention makes it well-suited for the task and the explanation I provide relies on intention’s characteristic features.

The problem for participatory intention arises from reflection on what it would be for an individual intention to settle a collection of agents on a course of action in the first place, given that an agent can only settle on something that they are in position to settle.\textsuperscript{6} If “What are we going to do?” is a single question, it is hard to see how each of us can settle it by way of our participatory intentions. If one person’s intention settles it, it looks as if no other person’s could. Yet, on the assumptions of this paper, each of us having a participatory intention directed towards our activity—that is, a state that settles our activity—is exactly what we need if each of us is going to be psychologically committed in the appropriate way to what we’re going to do. Given certain power relations, one person may be able to settle the issue for another, but, intuitively, collectively intending is not one person intending for all; it is all deciding or being decided. So, how could each participatory intention settle what we are going to do, without thereby settling it for the others and problematically bypassing their agency? Further, how could an agent regard their own intention as settling the issue without taking a demeaning view of the agency of their partners?

\textsuperscript{3} Exactly what the agent has to think about this resolution is a point of much contention.
\textsuperscript{4} See (Bratman, 2014, 16-18).
\textsuperscript{5} This methodology was developed in a series of papers by Michael Bratman, culminating in his Bratman (2014), where he calls it “constructivism”. The same limitations are explicitly acknowledged by (Bratman, 1999, 144). (Gilbert, 2000, 28-30) argues that interpersonal normative states actually do a better job of fulfilling these roles. Asarnow (2020) and Salomone-Sehr (2022) extend this methodology by attempting to construct forms of shared agency using individual psychological states other than intention.
\textsuperscript{6} This is known as the “Settle Problem” in the literature, and was originally raised by Velleman (1997).
A solution, proposed by David Velleman, is to appeal to conditional intentions. On this proposal, two (or more) people collectively intend if each has a participatory intention that is conditional on the other participant’s like intention (or participants’ like intentions). That is, each has an intention expressible by a sentence of the form “I intend to \( \phi \), so long as you intend likewise”. The first thing to note is that, in addition to being individual and conditional, the intentions involved are reciprocal—the condition on each is the matching intention of the other(s). The idea is that, assuming common knowledge of each others’ attitudes among the relevant individuals, when each person has a participatory intention like this, the condition on each is satisfied. Then, each is in a state that, taken together with the others, settles them on a course of action and represents itself as so doing. So, we have a collective intention to \( \phi \). Call this framework the “conditional intention account” of collective intention.

The most important question about this account is whether it plays the characteristic roles of intention across people. If it does, it’s a kind of collective intention, one that a full theory of sociality will include. I address this question in §4. There are, in addition, two further motivations. The first is the route to a solution it offers for the problem of how we genuinely share agency without any one of us imposing our own will on our fellow participants. A conditional intention doesn’t presuppose that it has discretion over what the others do, since it presupposes only conditional discretion. Because the condition on that discretion incorporates the agency of our partners, our deliberative question is settled without any individual settling the question for another. And because each agent only regards their intention as conditionally settling the matter, they do not need to be insensitive to the agency of their partners. Nor is there any issue with getting started. What a participatory intention settles on its own is up to an individual agent. So, there is nothing unreasonable about forming a solitary participatory intention when the agent reasonably believes the other will form theirs before the time for action arrives. Our jointly performing the action is then settled by the participatory intentions taken together. The second is that it commits us neither to any new individual psychological states nor to collective psychological states. Other attempts to construct something that plays the roles of intention for a collective often introduce new furniture to the world, for example novel modes of intending or novel interpersonal normative relations. For those who prefer minimalism, it is an advantage that this account introduces no new theoretical posits. Even a complete skeptic about collective intention needs to employ individual conditional intentions, since, to paraphrase Dorothy Edgington, a view of individual agency would be an idle, ineffectual affair without them.

---

7 The proposal is introduced in Velleman (1997) and further discussed in Velleman (2001, 2013). The possibility of using conditional intentions to give an account of collective intention is also discussed in (Bacharach, 2006, 139-141) and (Ferrero, 2009, 731-2, fn. 4).

8 For the rest of the paper, I assume that the participatory intentions of each are common knowledge among the participants. Common knowledge in one form or another is a standard assumption in the literature. See, for example, Bratman (2014), Gilbert (2013), and Tuomela (2005). It has, however, recently been challenged. See Blomberg (2016) and Ludwig (2016).

9 These two advantages are shared by the prominent account of collective intention given in Bratman (2014). However, as Bratman there acknowledges, and as mentioned in fn., Bratman provides only sufficient conditions, leaving open the possibility that there are other psychological structures that constitute collective intention. To my mind, this is the correct approach. It is probable that more than one combination of attitudes plays the roles of collective intention as defined here. A further background assumption of this paper, then, is that we should aim for an an account of each of these structures.

10 For example, Tuomela (2005).

11 This approach is used in Gilbert (2013).

12 Her claim is about the ability to think conditional thoughts more generally (Edgington, 1995, 235). It is in line with Ferrero (2009), who argues that the “deep structure” of almost all individual intentions is conditional.
2  INTERNAL, NON-PRECAUTIONARY, ELECTIVE, RECIPROCAL CONDITIONAL INTENTIONS

This outline of the account leaves open some important questions about the nature of these conditional intentions, a topic made more difficult by the way that the logic of conditional intentions differs from the logic of conditional statements. In response to a series of challenges based on that logic, I argue in this section that participatory intentions need to be a certain kind of conditional intention, namely the kind where the condition is reciprocal, internal, non-precautionary, and, in central cases, elective.

There are two forms of conditional intentions. Externally conditional intentions are conditioned on background features of the world that must be satisfied for an intention to exist in the first place. They are expressible by sentences of the form “If C, I intend to Ø.” Expressions of such intentions are predictions about what one will intend under some condition or will not intend if the condition fails to hold, but the condition doesn’t modify what the agent intends. Internally conditional intentions do qualify what the agent intends. They are expressible by statements of the form “I intend to (Ø if C)” and their existence is not dependent on the obtaining of the condition, nor the agent’s awareness of that fact. It is just that, even if there is such a token intention, it doesn’t require of the agent who has one that they perform Ø unless they are aware that C is satisfied. So, what kind of conditional intentions make up collective intentions?

According to Abraham Roth, we have a dilemma that calls into question whether the participatory intentions we are looking for could be reciprocal conditional intentions at all. He argues that if the condition is external and reciprocal it fails to settle anything. The thought is that if my intention is externally conditioned on yours, and yours on mine, then the existence of my intention is conditioned on itself. If I intend to Ø, you intend to Ø, and if you intend to Ø, I intend to Ø. So, I intend to Ø, if I so intend. But this doesn’t close the deliberative question: am I going to Ø or not? It is still up in the air. Since this doesn’t settle the agent on any course of action, it is best seen as not an intention at all, strictly speaking. Whatever they are, states like this are not candidate participatory intentions. This argument successfully rules out the possibility of reciprocal, external conditional intentions.

But it does not apply directly to internally conditional intentions. Roth mounts two separate arguments against them. First, he argues that internal, reciprocal intentions are not interdependent in the right way. Genuine interdependence, according to Roth, requires that one participant doesn’t have the intention if the other doesn’t. Internally conditional intentions are then independent, since I can’t remove your intention simply by dropping mine, as I could with externally conditional intentions. This is a problem because normally, “I wouldn’t have an intention concerning our walk unless you also had such an intention” and “it is the recognition of this dependency that motivates the conditional approach in the first place” (Roth, 2014, 630).

The answer to this challenge is that it mischaracterizes what is important about the interdependence involved in reciprocal conditional intentions. What is important about that interdependence is not the origin or existence of token intentions in the minds of the participants, but

---

14 For discussion about the distinction between internally and externally conditional intentions, see Cartwright (1990).
15 I’ll call the condition expressed by the subordinate clause of statements of the form “antecedent” and the other part “the consequent”.
16 The full discussion I engage with here is (Roth, 2014, 629-31). See (Roth, 2004, 373-80) for an earlier discussion.
the activation of these intentions, since it is this activation that characterizes the rational and causal roles of an intention in relation to the action or state of affairs specified in its content.

Before I expand on this point, let me note some features of activated conditional intentions important for the following discussion. First, a conditional intention is \textit{activated} when the condition is satisfied and the agent takes that to be the case. Second, \textit{unactivated} conditional intentions are not without force. As intentions, they still make rational demands on the agents who have them, given the norms of intention spelled out in §1, and tend to lead agents to behave in characteristic ways. For example, an agent with an unactivated conditional intention is rationally required and tends not to behave in ways that would make performing the action impossible if the condition obtains. Third, activation marks a transition point of a single attitude, not the formation of a new attitude or the transformation of a conditional attitude to a categorical one. Consider what happens when an agent with what they think is an activated conditional intention discovers that the condition is not satisfied. That attitude is not governed by the rational requirements on a categorical intention, but rather those of an unactivated conditional intention, which, at the outset, include assessing the likelihood of the condition obtaining.

Participatory intentions are activation interdependent in several ways. First, the intentions are logically and rationally interdependent. Each person’s intention plays a particular role in the other person’s practical reasoning. Your intention is required to satisfy the condition on my intention, and mine is required for yours. And then, once I take it that the condition on my intention is satisfied, that is once I believe you also have a participatory intention, I am under a host of newly specified rational pressures concerning the action. Our participatory intentions are thus interdependent both in their ability to satisfy each other’s antecedents, and in determining the specific demands of rationality we are under.

They are also causally interdependent. Roth frames the causal dependence of genuine interdependence as backward-looking, that is, as what causes each intention. He assumes that my intention has to cause yours or yours has to cause mine in order for them to be interdependent. But this is not the only way for participatory intentions to be causally interdependent. Their causal interdependence may be \textit{forward-looking}, that is, about what these intentions cause and represent themselves as causing rather than what causes them. And internal, reciprocal conditional intentions are causally interdependent in this way. My intention represents itself as having causal power in the direction of the action or outcome it specifies only in the presence of yours, and vice versa. Neither represents itself as causing us to act on its own. But when co-present, commonly-known, and effective, they do tend to cause us to act. The conjunction of these logical, rational, and causal features spells out the sense in which each intention settles the question of what we are going to do only conditionally on the other intentions. It also captures a sense of interdependence that does justice to what we think is involved in the relation between partners acting together, and it does so better than Roth’s proposal about what genuine interdependence amounts to. The relevant issue

\footnote{These points follow and are developed further in (Ferrero, 2009, 707-710). See also the debate between Ferrero (2015) and Ludwig (2015) for an elaboration and defense of these points in response to an alternate understanding of conditional intentions.}

\footnote{Of course, if I falsely believe that you have a participatory intention, I may begin to take steps that I think will lead to our \(\phi\)-ing. That is, I may start to perform the actions that would constitute my part in our \(\phi\)-ing. But that is not what the intention represents itself as causing. The content of the intention is not to act when one believes (perhaps falsely) that the other had the intention, but to act when the other actually does. Further, as I discuss in the next section, it represents itself as causing \textit{our} performance of the action, not my performance of what would be my parts. And it does so in a way that depends on the other having the corresponding participatory intention. Therefore, it does in fact tend to cause our \textit{performance} only when the other intention is present.}
between partners is not normally whether one’s intention exists without the other or is caused by the other, but the way in which each intention is activated by the other.

Roth’s second challenge to internally conditional reciprocal intentions is that, under the assumption of common knowledge, they may not commit us to any course of action either. Roth takes it that intentions with conditional contents give their agents two rationally permissible options. One can either go ahead and act in ways appropriate to \( \phi \)-ing or, for cases in which one has the requisite control, one can see to it that \( C \) is not satisfied. Because of the common knowledge assumption and the reciprocity of the intentions, each participant has this control. Seeing to it that \( C \) is not satisfied in our case means seeing to it that one’s potential partner doesn’t have a participatory intention. And one can ensure that is the case by dropping one’s own intention and making it clear one will not form a new one. Because these intentions are internal, the other’s intention doesn’t simply go away. But since there is little point in retaining a conditional intention if one knows that the condition won’t be satisfied, it makes sense for the other to drop their own intention. Once the other’s intention has been dropped, the dropping of one’s own intention makes sense as well. In making it so that the antecedent on my conditional intention doesn’t and won’t obtain, I’ve done nothing that goes against what that intention committed me to doing, according to Roth.  

On one way of understanding this line of argument, it is about what it takes to satisfy a conditional intention. If the content of a conditional intention were a material conditional, satisfying such an intention would be as simple as falsifying the antecedent. On this assumption, we get a strong version of the argument above. It is not only that one has not acted against one’s own intention by making it so that one’s partner doesn’t have their participatory intention, one has actually satisfied it by dropping it—a result that would recommend rejecting the conditional intention account. But this assumption is false. Carrying out a conditional intention is not the same as making a material conditional true, whether the agent has control over \( C \) or not.  

Conditional intentions are formed because an agent aims at \( \phi \)-ing when \( C \) obtains. That is their purpose in practical reasoning, that is why agents like us adopt them, and that is what makes them so important to understanding agency in general. When \( C \) can no longer obtain, it is not as if an agent succeeds at what they are conditionally pursuing. Succeeding at making \( C \) no longer obtain is succeeding only at depriving the intention of its point. A nurse who has an internally conditional intention expressible by “If the patient is still alive tomorrow, I will change his drip” cannot carry out their intention by killing the patient. A conditional intention is not satisfied when its condition can no longer obtain; it simply becomes moot. It is only carried out when \( C \) obtains and the action is executed. This understanding of the argument then does not give us a reason to reject

---

19 In the body of the text, Roth assumes his notion of interdependence while making a version this argument (Roth, 2014, 631), but suggests in a footnote that this assumption isn’t necessary and gives a slightly different version of it without that assumption (Roth, 2014, 649, fn. 19). Since I have given a different picture of what the interdependence of internally conditional reciprocal intentions amounts to, I focus on Roth’s footnote version here.

20 See (Ferrero, 2009, 703-6) and (Ludwig, 2016, 58) for discussion of the general issue.

21 This example is modified from (Edgington, 1995, 290).

22 The term “moot” is used this way in Ferrero (2009). Note that this is often also claimed of conditional imperatives. Vranas (2008), for example, holds that a conditional imperative is satisfied when both conjuncts are satisfied, violated when the antecedent is satisfied and the consequent isn’t, and avoided when the antecedent isn’t satisfied. He also provides a list of references to people who hold similar views about conditional imperatives (Vranas, 2008, 553, fn. 10).
internal, reciprocal conditional intentions, it instead gives us another reason not to interpret the content of conditional intentions as a material conditional.\(^\text{23}\)

Roth’s version of the argument, however, is more careful. He does not say that one satisfies their participatory intention by dropping it, but rather that one does not violate any rational requirements in so doing. For some conditional intentions, this is true, namely for “precautionary plans”.\(^\text{24}\) Say I intend to protest if another act of violence occurs. That intention places me under rational demands, for example to do not things that would make it impossible for me to protest if violence occurs. Yet, I may certainly also try to make it so that no such violence does occur. My intention is a plan I would much rather not carry out, but one which, should the undesired condition obtain, I would prefer to have. Nonetheless, it is still the case that if I succeed in avoiding violence, I haven’t carried out my conditional intention, I’ve simply succeeded in making it moot. Note that this example only makes sense because there is some question about whether I can avoid violence. If the occurrence of violence were under my direct and immediate control, there would be no need for this intention. I could simply make sure no violence occurred. This conclusion is general. Precautionary plans where an agent has direct and immediate control over the condition are pointless.

And, as we have seen, we have some control over the participatory intentions of our partners because they are reciprocal. I may not be able to satisfy my own participatory intention by dropping it, but I can make it moot by doing so. Because the intention is moot if dropped, dropping it doesn’t violate any rational requirements. Intuitively though, there is a crucial difference between my partner’s corresponding participatory intention and violence occurring. If we’re partners, your corresponding participatory intention is not an undesired circumstance with which I have to deal, as it appears to be in Roth’s challenge; it is rather exactly what I am hoping for in seeing you as aiming likewise to act together with me. In answer to Roth’s argument then, we should specify that the conditional intentions that constitute collective intentions incorporate this feature of our attitude towards our partners and towards \(\phi\)-ing, and so are non-precautionary plans. We can understand non-precautionary intentions as a subset of internal conditional intentions involving a further commitment on the part of the agent not to exercise whatever control they have over \(C\) by endeavouring to make it not obtain.\(^\text{25}\) And, we can understand participatory intentions as non-precautionary, internal, reciprocal, conditional intentions.

---

23 Edgington suggests non-truth-functional analyses of conditional belief, expressions of desire, and promises for reasons similar to those appealed to here in the case of conditional intentions and in the previous footnote in the case of conditional imperatives (Edgington, 1995, 288).

24 See (Ferrero, 2009, 706) for this label.

25 Because the nature of shared agency includes, intuitively, aiming to do something together with another, I suspect that this commitment is usually in the background of an agent’s practical reasoning. Intending to moot an intention one has by falsifying the antecedent is directly counter to the intuitively specified aim. There are two further relevant points. First, this commitment leaves open regular reconsideration one’s entire intention, consistent with the norm of stability, since that doesn’t include an intention to moot one’s own conditional intention. Second, to the extent one has control over the obtaining of an antecedent, one may usually exercise it. However, even in this case, agents often place additional demands on how that obtaining comes to be. It may go against how one conceives of an instance of shared agency that the other’s intention is formed on the basis of a promised monetary reward, for example. At times, this may also involve further commitments on the part of the agent. For discussion of possible restrictions on making a condition obtain, see (Ferrero, 2009, 737, fn. 54).
There is one more distinction between kinds of internal conditions that may be relevant for genuine interdependence. Internal conditions on intentions may either be enabling or elective. An enabling condition is something that needs to be in place for the action to be possible; that is, it concerns the conditions under which an agent can perform an action. For example, I may intend to work at the library if it is open. An elective condition, by contrast, is a restriction that an agent places on what it takes to succeed in executing an action, usually based on conditions under which they see it as worth doing. For example, I may intend to work at the library if my neighbour is practicing drums. Elective conditions therefore engage the intender’s agency in a very different way than enabling conditions. A reason for thinking that participatory intentions should involve elective conditions is that genuinely acting together may require that agents choose to restrict the satisfaction of their own participatory intention to situations that involve their partners’ intentional participation. And, in standard cases at least, this will certainly be so. However, those who think that acting together is consistent with participants seeing their partners’ participation as simply a matter of background necessity for completing some task may disagree. We may therefore prefer to treat this distinction as revealing two different kinds of relation between participants, one based on the desired participation of the others and one based on the required participation of the others. A more careful specification of this relation between partners is the topic of the next section.

So, reciprocal, conditional intentions expressible by statements of the form “I intend to \( \phi \), if you intend likewise” are candidate participatory intentions, as long as their conditions are internal and they aren’t precautionary plans, because conditional intentions with these features are genuine commitments and involve genuine interdependence, contra Roth.

3 | SHARED AGENCY AND THE ROLES OF COLLECTIVE INTENTION

An important test of a theory of collective intention is whether it explains a particular kind of social phenomenon—shared agency. Specifying this kind of social phenomenon is a way of translating the kinds of individual behaviours intentions explain to the collective case. An initial look at the structure of the conditional intentions we have so far will suggest they aren’t able to play the translated roles. Responding to this worry suggests slightly modifying their content.

Consider the following contrast between two cases. In the first case, we have commodity traders who behave according to a simplified version of an economic story about how prices in a market are determined. Roughly, the traders buy and sell for their own purposes based on extensive

---

26 Ferrero (2009) calls these two kinds “preconditions” and “restrictive” conditions. Ludwig (2016) calls them “enabling” and, following Bratman (1985), “reason-providing” conditions. The label “enabling” condition for the second kind is also used in Klass (2009).

27 More formally, following Ludwig (2015), we can say that \( C \) is an enabling condition on \( x \)’s \( \phi \)-ing iff \( x \) cannot \( \phi \) unless \( C \) as a matter of aletic or nomic necessity. So, for the example to work, I must be thinking of working in the library in a way that precludes things like breaking in. Davidson (2001) famously denied that these are genuine conditional intentions; most disagree. Further distinctions between kinds of enabling conditions are discussed in (Ferrero, 2009, 702).

28 Presumably for this reason, Velleman assumes that the conditional intentions involved in collective intention are elective (Velleman, 1997, 46, fn. 27).

29 Compare a parallel distinction drawn by (Bratman, 2014, 70-72) between “desirability-based” and “feasibility-based” persistence interdependence.
knowledge about each other’s histories, contextual factors, trends, and so on. This process may involve complex interpersonal strategic reasoning, for example predictions about other traders based on mutual mindreading, and it may involve mutual responsiveness, so that changes in the behaviour of one trader reliably result in changes in another trader’s behaviour. The outcome of these individual actions of buying and selling is the price. This is plausibly seen as a collective outcome, of some kind, since setting the price cannot be understood distributively. It is not true of each trader that they set the price. The price is produced by way of the non-accidentally-related contributions of each. Yet at no point is that particular price something the traders aim at bringing about together; it just happens. In the second case, imagine that all of the individual behaviours are the same—the same commodities are sold for the same amounts—but they’re undertaken as a result of a plan held by a group of traders to keep the price at a particular level by executing exactly that pattern of trades. That is, the traders are engaged in price-fixing.

Intuitively, the price-fixers are acting together in a way that the traders in the standard story aren’t. What our intuition is picking up on here is the way in which the people who generate the price are moved by it and each other. First, the agents in the price-fixing case coordinate their actions in a way that tracks their shared goal, monitoring the others behaviour and adjusting where appropriate. Second, they’re disposed to support and help their co-conspirators in executing their parts and to engage in the further deliberation about how to reach their goal. And third, there is a background framework for bargaining and resolving conflicts in preferences about how the action is to be carried out. That’s not the case for the regular traders. For them, the resultant price doesn’t need to feature in their individual reasoning or conception of what they’re doing, nor in the way they conceive of their relation to one another. They aren’t disposed to do what they can to ensure that some particular price comes about by engaging in deliberation or bargaining with each other. There are therefore three features that distinguish the shared agency of the price-fixers from the collective-outcome-producing traders: (i) interpersonal coordination of action tracking a shared goal, (ii) dispositions to support and deliberate about execution, and (iii) bargaining against a background framework for resolving disagreement.30

The difference between the traders and the price-fixers is important to us. Whatever we think about the morality of markets, the price-fixers are doing something wrong in a way that the traders in the standard story are not. It makes sense to see this difference in terms of the intentions of the participants with respect to the activity, which is the source of the broad agreement that appealing to collective intention is a natural way to explain it. A theory of collective intention needs to explain how these three features of shared agency arise between the participants, thereby showing how collective intention plays intention’s characteristic unifying roles for a collection of agents.

It might seem that the conditional intention account fails to capture this difference. After all, the thing a person with a participatory intention actually intends to do once the condition is satisfied is only their own action. But what the participants in shared agency aim at doing is something together, not at doing something on their own conditional on someone else doing the same thing. If all of them did only aim at their own action conditional on someone else’s intention, it seems like all we would have is parallel individual intentions that are limited to certain circumstances, and what we need for participatory intentions are intentions about our activity. Conditional intentions of the kind we have so far specified, one may think, will not result in a

---

30 This characterization of shared agency follows Bratman (2014). Even Gilbert, who denies the “Intention Thesis” (see footnote) accepts that explaining shared agency so characterized is a mark in favour of a theory of collective intention (Gilbert, 2009, 195).
state that explains the tracking and adjustment required for interpersonal coordination, much less the other two features.

The way to allay this worry is to modify the content of participatory intentions so that they are explicitly about what we are going to do. Fortunately, intentions with contents like this are familiar from many accounts of collective intention. All they require is that we change the content from “to $\phi$” to “that we $\phi$”, where this is to be understood distributively. That is, it is an intention that we both $\phi$ in the way in which two agents might $\phi$ in parallel, rather than an intention to engage in shared agency itself, which would be circular. This makes it clear that each agent not only intends their own $\phi$-ing, but the others’ $\phi$-ing as well. Some have objected that intentions of this form involve an undesirable bypassing of others’ agency, but, as we have seen, here the contents are embedded within a conditional intention where the condition goes by way of the other’s agency.

Participatory intentions are then expressible by a statement of the form “I intend that we $\phi$, if you intend likewise”, where, in paradigm cases, the condition is internal, non-precautionary, reciprocal, and elective. With this and the outline from §1, we have a full statement of the conditional intention account. And we have a clear view of the features of shared agency this psychological structure is to explain.

4 | EXPLAINING SHARED AGENCY

The issue now is whether, and if so how, this structure provides that explanation. The first thing to clarify is how this structure of intentions commits us to the action in the content of each. Working through this issue highlights a property of these collective intentions that follows from the properties of participatory intentions discussed in §2 and is crucial to their ability to explain the normativity—rational and interpersonal—of shared agency.

The problem is that it appears that the conditions on each intention aren’t satisfied. For simplicity, I’ll consider a two-person case. Take the general form, “I intend that (we $\phi$ if $C$)” and say that $C$ is just “you intend that we $\phi$”. Then, since what is required to satisfy that condition on the first intention is the categorical intention of the second person, the second person’s conditional intention doesn’t satisfy it. Because each of the individuals has a conditional intention rather than a categorical one, none of the antecedents are satisfied. One available response is to change the conditions so that they only require the activated conditional intentions of the others. So, now “$C$” is “you intend that we $\phi$ if I intend that we $\phi$”. But this fails to solve the issue because what is required to satisfy the condition on each is an intention that is conditional on the activated intention of the other, and that is not the intention that each of us has according to this suggestion. It looks like the only way to solve this issue is to give up on reciprocity. Say the first person has the intention expressible by “I intend that (we $\phi$, if you intend that we $\phi$, if I intend that we $\phi$)” and the second person has “I intend that (we $\phi$, if you intend that we $\phi$)”. Then the second person’s intention satisfies the condition on first person’s intention. In turn, the first person’s activated intention satisfies the condition on the second person’s intention. This solution preserves

31 (Velleman, 2001, 121) argues that the original formulation implicitly leads us to the modification suggested here. Whether or not that argument is sound, I think we should make that modification.

32 They are introduced in the context of shared agency in Bratman (1992).

33 The circularity worry is raised in Peterssen (2007) and responded to in Bratman (2014), to my mind successfully.

34 This worry is raised in (Gilbert, 2013, 44).
the conditionality of each person’s intention on the other. But, because there is always one person who intends with one less epicycle on the condition, it violates reciprocity.\textsuperscript{35}

The original statement of the account contains a response. Looking at the proposal more closely, the condition is not that the other intend, but that the other intend likewise. I argue shortly that the antecedents on intentions like that are satisfied when each has an intention of the same kind. But first, the cost of this framing of the intentions, which Velleman considers and accepts, is that the content of the attitudes is “gappy”, in the sense that it is not finitely completable.\textsuperscript{36} Trying to specify the content of the first intention requires specifying what would count as a “like” intention on the second person’s part. Doing that, however, leads us to the “likewise” in the second intention, and, in turn, to the problem of specifying what a “like” intention on the first person’s part would be—that is, back where we started. Each of these intentions is then self-referential in a potentially problematic way, since trying to fill out the contents leads to an infinite regress. Since we can never finish the specification, the contents of each are incomplete or indeterminate.

The central reason Velleman thinks we should not balk at accepting attitudes that have gappy propositions as their contents is that it doesn’t require us to accept indeterminate truth conditions for such contents. Each person is aiming to frame intentions similar to the other person’s, and if we take “likewise” to mean “determinate and indeterminate in the same way, if any”, when they frame participatory intentions as proposed here, they succeed. They succeed since “it is perfectly determinate whether either person’s intention has the same potential indeterminacies as the other’s” (Velleman, 1997, 45, fn. 26). That is, the intentions in our case do have the same determinacy and indeterminacy. The determinate part of the content is the collective behaviour, which is the same in both, and the indeterminate part is the self-referentiality in the antecedent, which is the same in both. Like a ratio that can be set without a determinate numerator or denominator, they are similar in the way each requires. So, when each of us has this kind of participatory intention, the conditions are satisfied and each of us is committed to our activity.

People with intentions like these are deeply interdependent, moreso than has been noticed before. And it is this radical interdependence that explains how this structure fulfills the characteristic roles of collective intention and is capable of serving as a basis for interpersonal obligation. Say we have a collective intention according to this view. It is not just that I intend conditional on your intending, I intend conditional on your intending conditional on my intending, and so on. And you intend conditional on my intending, conditional on your intending, and so on.\textsuperscript{37} The conditionality of my intention on yours is in the condition on your intention, and the same for

\textsuperscript{35} Goodin (2012) considers a solution like this in a different context, namely excuses available to an agent based on other’s behaviour and attitudes. Since he is not concerned with collective intention and reciprocity, he accepts it.

\textsuperscript{36} Velleman’s discussions of the issue are at (Velleman, 1997, 45-6, especially fn. 26), (Velleman, 2013, 108-9) and (Velleman, 2015, 235-6, especially fn. 35). In the context of a discussion of exchange, Julius (2013) proposes different intentions with self-referential contents like this. He ends up rejecting them, not for the implausibility of such contents, but for their inadequacy as an account of exchange.

\textsuperscript{37} This is a complex attitude for a human being to possess, so complex in fact that one may doubt whether a real human being could hold it. There are, however, some considerations that I believe should temper skepticism. First, we seem in general to be able make sense of such contents. Take an example of Velleman’s from a British publisher who would place the following notice on the copyright page: “This book is sold subject to the condition that it shall not be lent, re-sold, hired out, or otherwise circulated without the publisher’s prior consent and without a similar condition being imposed on the subsequent purchaser” (Velleman, 2013, 108). We can grasp what this sentence expresses without the content being finitely completable. Second, in my experience, deciding to do things with others in ways that employ interdependent attitudes of this kind is quite common. Faced with an intriguing possibility, people often turn to one another with an attitude along the lines of “I will if you will”, and such attitudes often result in action. It is not usual to articulate that attitude along the lines of the construal I’ve offered. But one of the advantages of the construal is that it explains how these attitudes
your intention on mine in my intention. When this kind of interdependence exists between us, we both know that what each intends, in a sense, determines what the other intends. 38 You have made plans conditional on me, and I have framed an intention that is activated only when you have an intention that is conditional on what I intend. Thus, I am taking my cue from you, just as you are taking your cue from me. This activation interdependence between participatory intentions provides an explanation of the pre-theoretic idea that acting together involves aiming to do something with another, while taking them to aim at doing that thing with you.

It also accords with the philosophical idea that collective intentions involve particular dispositions to be flexible about roles, support other participants, and persist through difficulties, dispositions that rule out certain cases of deception, deceit, or instability, discussed below. That is, the idea that collective intention explains shared agency. Traders in a market may have intentions that are conditional on each other, but they don’t have intentions conditional on the others’ intentions being conditional on their own.

As we have seen, when each of us has a participatory intention, we also have an activated intention towards our activity, which means that each not only intends their own actions in accordance with $\phi$-ing but the corresponding actions of the others as well. As such, we each are psychologically committed to what we’re doing in a way that is constrained by the characteristic rationality requirements on intention. This puts each of us under a host of rational pressures. From agglomerativity, I can’t rationally form an intention to do something else that would make our $\phi$-ing impossible. From stability, I can’t rationally drop my intention for no good reason. Also, there are limitations on the conditions under which reconsideration is rationally permissible, which makes our participatory intentions durable. Nor is either of us less committed to our $\phi$-ing because our participatory intentions are conditional. 39 It is true that each of us is committed to our $\phi$-ing under a smaller set of circumstances, those where the other is committed to our $\phi$-ing too, but that doesn’t entail anything about how committed we are when those circumstances obtain. Since our activity is what each of us intends, we are durably disposed and rationally required to act in ways that promote our achieving our aim. This structure therefore plays the first translated role of intention. We are set to undertake interpersonal coordination of behaviour towards a shared goal.

But each of us having an intention towards our activity is not enough, since “our activity” in the content of a participatory intention is neutral with respect to its own involvement of collective intention, and so neutral with respect to the intentions of the other participants. Because of this, it is possible to have an intention expressible by “I intend that we $\phi$”, while intending that we $\phi$ by means that do not include the participatory intentions of the other. 40 Each of us having intentions like this doesn’t ensure that we are really engaged in shared agency, even if we end up $\phi$-ing, since we won’t be disposed to support, negotiate, and deliberate in the appropriate ways, leaving as rationally permissible some forms of deceit and manipulation inconsistent with genuinely sharing agency. 41 Explaining these further features then requires turning to the condition on each of our intentions. Since the condition is internal, it is not simply a prediction about what you intend; rather, my participatory intention is only carried out when our activity is executed

work—how they get us to act—despite their complexity. Each only needs to be satisfied that the other has an intention that mirrors their own; they don’t need to pursue an attempt to complete the specification of the content.

38 What it determines is whether the condition is satisfied, not the content simpliciter.

39 Yaffe (2004) considers “degrees of commitment” with respect to conditional intention.

40 These are the so-called “Mafia Cases”, discussed in, for example, (Bratman, 2014, 49).

41 Some conflict between participants is still possible in shared agency, which leads Bratman (1992) to develop a more collaborative form of sociality he calls “Shared Cooperative Activity”.


and \( C \) obtains. The success of your corresponding participatory intention is a success condition on my own intention. And since it is non-precautionary, I am committed to not mooting it.

Because my intention can’t be carried out without the success of yours, all the same dispositions and rational pressures apply to me with respect to to your participatory intention. That is, I am required to do what I can to ensure our activity goes by way of your intention. In general, this involves supporting the maintenance of your intention by working out how to carry out the action in line with how you see your intentional participation. Not only does this properly rule out cases in which one agent tries to bypass the intentional participation of the others, it also motivates us to engage in deliberation about roles and provides a background for bargaining.

The interdependence between the conditions on our participatory intentions forces us to coordinate our planning in order for them to be successful. Neither of us is in a position to pursue ways to execute our activity that subvert the other’s intentions about how to carry the action out. Why? If I fail to behave in a way that takes your intention into consideration, it signals that I don’t actually have the intention that would satisfy the condition on your intention. Rather it signals I have an unconditional intention that we \( \phi \), and am using you to do so. But if that is the case, since the condition on your intention isn’t satisfied, your intention is moot. And because there is little point in maintaining a moot conditional intention, you are likely to drop it. Then, my intention is moot too. On pain of mooting my own non-precautionary intention, something I am committed not to do, I am motivated and rationally required to deliberate and negotiate with you about how to carry out our activity on terms that preserve and support the success of your intention. Further, because I believe that your intention is necessary to my intended end, I am committed by means-end coherence either to monitoring your intention and its progress or to dropping my intention. So, as long as I don’t drop my intention, I will engage in the kinds of mindreading monitoring your intention requires. And I will be responsive in my behaviour to cues about how our joint deliberation is proceeding. Since all of this is the same for you, the mindreading and responsiveness is mutual. We’re both set to work out what we need to do to execute our intentions and to resolve differences in preferences about that execution in ways that do not undermine either’s participatory intention. This structure therefore explains both (ii) the dispositions to support the other’s activities and plans and to partake in shared deliberation, and (iii) the background framework for bargaining.

5 | INTERDEPENDENCE AND INTERPERSONAL OBLIGATION

Some philosophers think that in addition to rational pressures, shared agency involves interpersonal obligation.\(^{43}\) For them, the conditional intention account faces a daunting challenge

---

\(^{42}\) This form of shared agency will be slightly different that the one emphasized in Bratman (2014), which requires a stronger form of mutual responsiveness in that it must be successful rather than only intended (Bratman, 2014, 78-82) and a further condition about meshing subplans, which may conflict with competitive cases, like playing chess (Bratman, 2014, 55-6). However, (Bratman, 2014, 55) also notes that, to an extent that depends on some further specifications of the theory, much of the value of meshing subplans for unifying agency across people is also present in requiring that the completion of the action goes by way of the others’ participatory intentions. The fundamental issue is that in so far as this structure of conditional intentions explains the three general features of shared agency—that is, in so far as agents with these states are set to engage in social interactions with those features—it plays the roles of collective intention.

\(^{43}\) For Gilbert (2013) and Roth (2014) these obligations are inherent in or a necessary part of shared agency. Alonso (2009) also argues for a tight connection between shared agency and interpersonal obligation, but not an inherent or necessary one.
explaining this fact, since conditional intentions aren’t the kinds of things that generate mutual, directed obligations. The interdependence of our participatory intentions provides an answer to this challenge as well.

The relation between interpersonal normativity and shared agency is highly contentious. One strategy for those who favour the conditional intention account would be to simply deny that there is such a close connection between them. However, that response is needlessly restrictive. Combined with a background normative theory, it is capable of generating a range of positions with respect to what we owe each other when we act together, including one that treats them as inherently related.

The aim of this section is modest. I do not develop an account of mutual obligation in shared agency. Instead, I attempt to show that even those who take a tight connection between shared agency and mutual obligation as a starting point should not dismiss the conditional intention account. Specifying how conditional intentions relate to mutual, directed obligations is not as unfeasible as some think once we take into consideration the kind of mutual recognition and interrelation required by activation interdependence.

Here are two examples of the challenge. Margaret Gilbert claims that it is plausible to suppose that a view of collective intention based on conditional intentions “can be ruled out a priori” because it is “hard to see how a combination of conditional personal commitments can, through satisfaction of their conditions, create a set of commitments such that the committed people are answerable to one another for violating them” (Gilbert, 2013, 45). Roth agrees, expressing the thought forcefully in the following passage:

If I were to form the conditional intention to go for a bike ride so long as it is sunny, it doesn’t follow that when it is sunny, I somehow owe it to the weather—that I have a ‘commitment-to-the-weather’—to go for a ride… It doesn’t seem to me that anything of significance would change were I to condition my bike ride on the behavior or intentions of other people. One’s conditional intention to take the surface streets if many people take the 405 does not generate in one any commitment to those many people on the 405 to take surface streets. Nothing intrinsic to the conditional intention itself generates any such commitment. So conditional intentions will not account for the contralateral commitments in shared agency (Roth, 2004, 374-5).

Roth is correct that the connection the conditional intention account posits between collective intention and interpersonal normativity shouldn’t entail commitments-to-the-weather. That would follow if, say, the connection relied on the claim that all individual conditional intentions generate obligations. So, what the conditional intention accounts needs, then, is a feature of the particular kinds of conditional intention that make up collective intention that could serve as a trigger or ground of interpersonal obligation. That ground or trigger would not thereby be present in other cases when agents form regular conditional intentions.

It is unsurprising that there are no interpersonal obligations in the weather case, since there is only one person. Roth’s point though is that this fact doesn’t change once we introduce other people. He is right about that as well. There are many instances in which an agent framing intentions conditional on the behaviour or intentions of others does not generate interpersonal obligations. This is true even though these kinds of cases do involve recognition of agency, since such a conditional intention presupposes that the other is an agent capable of forming intentions. And it is true if this recognition is mutual, say if the other person forms a conditional intention based on some intention of mine. So, we may conclude, with Roth, that the mutual recognition of agency
entailed by some structures of conditional intention doesn’t have the right profile to serve as a trigger or ground of interpersonal obligation.

But as we have seen, the conditional intention account posits a stronger interrelation between the attitudes of the participants. When we have a collective intention on this theory, it is not just that we mutually recognize each other’s ability to form intentions, and then make our intentions dependent on a prediction about what others intend; we go the further step of representing that mutual recognition in the content of our own intentions. This follows from the special kind of interdependence between these intentions. Each of us conditions our own intention not only on the intentions of others, but on the others’ intentions being conditional on our own too. This places our interdependence in the condition on the intention, and since these intentions are internal, this condition is in the content of each intention. In addition to the recognition of agency involved with the other examples, there is a further recognition of the other’s special role in satisfying the condition on my intention. That is, for us to have a collective intention on this view, I not only have to recognize you as an agent, I have to represent you as an agent who is at the same time recognizing my agency by making your intention conditional on mine. Since you’re doing the same, we’re making our intentions conditional on this mutual recognition of agency. None of this is the case for the prospective bike-riders, the 405 commuters, or the traders in the standard story.

Reserving my intention for cases in which you are depending on me to determine what you do and what specific rational demands you are under while you are doing the same is also different than mutually generating expectations. There is a sense in which by discharging the condition on your intention, if you have it, I am making up your mind too with respect to our action. I am transforming the rational and causal force of your intention, from one limited by having an unsatisfied condition to one directed at our action. All of this is the same for you with respect to my intention. What we do in framing such intentions is not primarily a matter of reliance or assurance, because it is not a matter of prediction, but rather a form of co-determination. What I intend determines what you are rationally required to do, and vice versa.

This difference between participatory interdependence and other situations in which we form conditional intentions is enough to provide an initial answer to Gilbert’s and Roth’s challenges. If shared agency involves interpersonal obligation, it may be triggered by or grounded in the specific kind of interdependence, the kind of practical co-determination, these collective intentions involve. There is therefore no need to claim that all internal, mutual-agency-recognizing conditional intentions generate obligation. Many questions about the normative principles that would ground or explain interpersonal obligation in shared agency remain open, of course. But they do not need to be answered to show that the conditional intention account cannot be ruled out a priori.

6 | CONCLUSION

Individual, reciprocal, and non-precautionary conditional intentions towards a collective activity are participatory intentions. If each of a number of people has a participatory intention towards the same activity, and that fact is commonly known amongst them, those people collectively intend to do that thing. Collective intentions understood this way do justice to the intuitive idea that when we aim at doing something together, we each see the others as likewise aiming at

44 Alonso (2009) proposes this is as the source of obligation in shared agency.
doing that thing with us. They also play the roles associated with collective intention in the philo-
osophical literature, thereby explaining shared agency. And they provide a basis for normative
 theorizing about the interpersonal obligations that may exist between people engaged in shared
agency. They do all of this because the structure of attitudes they involve creates a special kind of
interdependence between the participants.

ACKNOWLEDGMENTS
My deepest thanks to David Velleman for conversations on the topic and comments on the paper.
Thanks also to Franz Altner, Stephanie Collins, Luca Ferrero, Niels de Haan, Javier Gomez-Lavin,
Carol Gould, Carlos Núñez, Grace Paterson, Herlinde Pauer-Studer, Lee Rachar, Abe Roth, and
Jules Salomone-Sehr, as well as three referees from this journal and the audiences at the Rea-
soning and Normativity VIII Workshop in Abu Dhabi in 2020, the Mississippi State University
Colloquium in March 2022, and the 2022 APA Pacific in Vancouver for helpful comments.

REFERENCES
University Press.
Philosophy, 94, 315–326.
Philosophy of Donald Davidson. Cambridge: Blackwell.
Clarendon Press.
Littlefield.


**How to cite this article:** Rachar, M. (2023). Conditional Intentions and Shared Agency. *Nous*, 1–18. https://doi.org/10.1111/nous.12452