LUND UNIVERSITY

# Intentional cooperation and acting as part of a single body

Blomberg, Olle

*Published in:*
Mind & Language

*DOI:*
10.1111/mila.12274

2021

*Document Version:*
Early version, also known as pre-print

Link to publication

*Citation for published version (APA):*
Blomberg, O. (2021). Intentional cooperation and acting as part of a single body. *Mind & Language*, *36*(2), 264-284. https://doi.org/10.1111/mila.12274

*Total number of authors:*
1

**General rights**
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

**Take down policy**
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

# Intentional cooperation and acting as part of a single body

Olle Blomberg

Lund University & University of Gothenburg

olle.blomberg@gmail.com

Pre-print version of a paper forthcoming in *Mind & Language*.

Please only quote from the published version.

**Abstract:** According to some accounts, an individual participates in joint intentional cooperative action by virtue of conceiving of him- or herself and other participants as if they were parts of a single agent or body that performs the action. I argue that this *notional singularization move* fails if they act as if they were parts of a single agent. It can succeed, however, if the participants act as if to bring about the goal of a properly functioning single body in action of which they would be parts. This latter version of the move manages to capture the cooperative character of joint intentional cooperative action. It does this without requiring of participants that they act on higher-order interlocking intentions.

**Keywords:** Intentional cooperation, group-identification, joint commitment, shared cooperative intention, team reasoning, joint intentional action

**Word length:** 11,993 (Acknowledgements excluded)

# 1 Introduction

We intentionally coordinate our actions to achieve common goals. This allows us to, for example, play games, go for walks, hunt and carry heavy pieces of furniture together. In doing these things, we are not interacting strategically. If we cook dinner together, we are not merely each making half a dinner in parallel, expecting the other to make the other half. There is one intentional action in which we all participate, each performing actions that are components of one and the same larger action. To use a technical term, we can say that we have and act on a 'shared intention' to cook dinner if we do it together as a joint intentional action.

Paradigmatically, we carry out our joint intentional actions in an intentionally cooperative fashion. I will call such an action a *joint intentional cooperative action* (JICA). A JICA is the result, as I will put it, of a 'shared cooperative intention'. JICA is thus a species of joint intentional action and shared cooperative intention is a species of shared intention. Shared cooperative intentions allow us to achieve common goals with a remarkable efficiency—a fact that surely partly explains why humans have been such a thriving and resourceful species.

Exactly how the notion of cooperation involved in JICA should be prespecified is a difficult matter. I will assume that it minimally requires that the participants of a shared *cooperative* intention cannot rationally and intentionally physically force each other to carry it out in a certain way (Bratman, 1993, p. 110; Pacherie, 2013, p. 1822). If our walking together is appropriately guided and coordinated by our shared

cooperative intention, then I cannot engage in such brute coercion to make you follow my preferred path.

We humans have also organised ourselves into larger social entities such as corporations and political parties. In some respects, it is appropriate to treat these entities themselves as individual agents (List & Pettit, 2011). In one sense, such individual agency at the group level is a more complex phenomenon than JICA. But from a bird's eye view, group-level agency is a just another form of individual agency. It is no more puzzling than rational agency in general. By contrast, JICA involves several rational agents sharing discretion and control over a single action.

In this paper, I critically discuss various accounts that are supposed to make sense of JICA in terms of participants' thinking of themselves and their co-participants as if they together made up a single body or agent. The aim of the accounts is thus to make sense of people acting together by appealing to the *idea*, in the minds of the participants, that it is as if there were a group-level agent or body of which they were all parts. JICA would thus involve each party conceiving of the intended joint activity as the action of such an imagined single agent or body. Furthermore, on these accounts, not much else is supposed to be required for capturing the intentionally cooperative character of JICA. I refer to this theoretical strategy as the *notional singularization move*. The strategy promises to shed light on the complex (JICA) by means of something more elementary (individual group-level agency).

I will argue that the *agent version* of the move fails because parties who act to bring about the goal of an additional single agent may intentionally and rationally, and without reasoning from false assumptions, brutely coerce the others in ways that are

incompatible with JICA. They can do this despite thinking of themselves and the others as if they were parts of that agent. However, I will argue that a *body version* of the move succeeds in capturing the cooperative character of JICA. The main contribution of this paper is the identification of the notional singularization move, the distinction between its two versions, and my argument that only the body version succeeds.

In section 2, I characterize the target phenomenon of JICA and present Michael Bratman's (1992; 1993; 2014) influential way of capturing its intentionally cooperative character by appeal to interlocking intentions of the participants. I also mention a possible motivation for the notional singularization move, namely that having interlocking intentions may be too conceptually demanding. In section 3, I show how the notional singularization move is exemplified in Margaret Gilbert's (2002a; 2008) joint commitment-based account of shared cooperative intention, as well as in accounts based on group identification and team reasoning (e.g. Pacherie, 2011; 2013). In section 4, I argue that these accounts fail to capture the intentionally cooperative character of JICA if they rely on the agent version of the notional singularization move. However, I argue in section 5 that they succeed at this if they instead rely on a suitably qualified form of the body version of the move.

## 2 Joint intentional cooperative action

Joint intentional actions include competitive activities such as playing chess or boxing (Gilbert, 1989, p. 441; 2008, pp. 503-504; Bratman, 2014, pp. 55-56; Ludwig, 2016,

p. 227). Joint intentional action is arguably also compatible with participants engaging in “modest sabotage” of the contributions of co-participants (Ludwig, 2016, p. 253-254). However, I will focus on JICA, which doesn’t involve competition or such sabotage.

This is not an arbitrary choice. JICA is the normal case. A shared cooperative intention is what under ordinary circumstances reliably leads to the successful execution of a joint action. In the realm of joint action, JICA is the equivalent of a composite singular intentional action, an intentional action made up of several component actions performed for the sake of a common goal. Such component actions are coordinated in a cooperative fashion by the agent over time, as well as at a time (such as when someone is simultaneously moving several limbs and digits while playing guitar). Minimally, this means that the components cannot be coordinated through brute coercion. Odysseus’ tying himself to a mast to avoid being affected by the calls of the Sirens is incompatible with his exercising intentional agency across time. Hence, his tying himself to the mast and then later unsuccessfully struggling to get free would not constitute a larger intentional action, even if he were at all times acting on or trying to act on an intention to sail past the Island of the Sirens. Similarly, in joint intentional action the participants treat each other as partners rather than as mere tools or objects to be used to achieve their own goals. In JICA, but not in activities involving competition or modest sabotage, the participants furthermore cooperatively coordinate their intentions all the way down to action.

In what I call brute coercion, an agent either intends to "bypass the other's intentions/subplans" or simply does not care whether or not she bypasses them (Bratman, 1993, p. 110).[1] In such coercion, the agent treats the other as a tool or object. Brute coercion is incompatible with JICA. If I intend that we cook dinner by my dragging and handcuffing you to the broken oven door against your will, so that you can only contribute by keeping the oven door closed, then we do not have a shared cooperative intention to cook dinner together, even if you also have an intention that we cook dinner and perhaps even act on it when you struggle to get free.[2]

Is brute coercion compatible with the successful execution of a shared intention and joint intentional action as such (in contrast, that is, to the successful execution of a shared cooperative intention and JICA)? Arguably, it could be.[3] Suppose that, while we are cooking together, the oven door unexpectedly breaks open at a crucial moment. To stop the heat from escaping the oven, I quickly lift you up and shove you against the oven door so that it is kept shut. Afterwards, I can justify my action to you in light of our overarching joint project of cooking together, and you may then indeed

---

[1] This is akin to what Michael Bayles (1972) calls "occurrent coercion", where "physical force is directly applied to cause behavior in another person." (p. 17)

[2] On the cooperative character of JICA, see e.g. Pacherie (2013, pp. 1821–2), Bratman (2014, pp. 48, 55), Gold and Sugden (2007, p. 110) and Ludwig (2016, pp. 225–28).

[3] I am grateful to Margaret Gilbert for making this clear to me, as well as for supplying the riff on my cooking example, which I have paraphrased here.

be glad that I, in the heat of the moment, didn't care about whether or not I intentionally bypassed your subplans, even though you resisted and protested at the time. This brutely coercive cooking is not an example of JICA though, which is caused and coordinated by a shared cooperative intention.

*Non-brute* coercion is compatible not only with joint intentional action, but also with JICA: Threats that change the incentives of others can be issued prior to, or during, the stage at which the shared cooperative intention is formed (Gilbert, 2008, p. 499; Tuomela, 2007, pp. 15, 167; Bratman, 2014, pp. 37-38, 55, 101). Similarly, non-brute "coercive" manipulation of one's future incentives is compatible with composite singular intentional action: Odysseus could have exercised intentional agency across time even if he had requested Circe to threaten his future self to set fire to his home in Ithaca unless he managed to resist the calls of the Sirens.

To capture the intentionally cooperative character of a JICA, it is not sufficient to require that the participants' common goal is that *they* bring about a specific outcome—what one might call a "joint goal" or "group goal" (Pacherie 2013, pp. 1831, 1835). Not that such a goal can be held by a single individual. It need not be shared. Now, an influential way of ruling out cases of brute coercion is to require that each participant intends that they *J*—that they cook dinner, say—by way of their own and the other's intention that they *J* and associated subplans (Bratman, 2014, pp. 48-52). In Bratman's terminology, the participants must have "interlocking intentions"

that they *J* as well as intended meshing of subplans (subplans mesh if they are co-realizable).[4] This arguably captures the intentionally cooperative character of JICA.

Readers familiar with Bratman's account of shared agency may be somewhat puzzled at this point. I have claimed that the target of Bratman's account is JICA, or perhaps some subspecies of JICA. However, Bratman (2014) refers to the target phenomenon of his account simply as "shared intentional activity" and explicitly distinguishes it from "a moralized notion" of "shared *cooperative* activity" that excludes both brute and non-brute coercion (p. 38, emphasis in original). However, there is a sense in which even Bratman's "shared intentional activity" is intentionally cooperative: The participants are normally disposed to help each other and treat each other as intentional co-participants rather than merely as tools or objects (ibid., pp. 48, 56). Furthermore, Bratman takes the mesh of subplans that the participants intend to achieve to be such that it reaches all the way to action (ibid., p. 79). His account is also supposed to rule out brute coercion but, as we have seen, joint intentional action as such is arguably compatible with some degree of brute coercion between participants. I will therefore refer to the target phenomenon of Bratman's account as 'shared cooperative intention' rather than 'shared intention'.

While Bratman's account captures the intentionally cooperative character of JICA, it places high conceptual demands on participants. They must have higher-order intentions regarding their own and the others' first-order intentions. Given that young

---

[4] Each must also believe or assume that their *J*-ing is a single goal such that it is intended by each (Blomberg, 2016).

children who lack such higher-order intentions nevertheless engage in what appears to be JICAs, it is worth looking for alternative ways of capturing their intentionally cooperative character (Pacherie 2013).

At first glance, the notional singularization move looks like an obvious alternative. A shared cooperative intention to cook together would be explicated in terms of each participant conceiving of their cooking together as if they were parts of a single agent or body that was cooking. There would be no appeal to complex higher-order intentions. But how is the move supposed to capture the intentionally cooperative character of JICA? My aim is to answer this question, independently of whether or not the move will help capture what is going on in the joint activities of young children. In the next section I present two specific examples of the notional singularization move in some detail. In sections 4 and 5 I then argue that the move only succeeds when given a specific interpretation.

## 3 Two examples of notional singularization

I will describe two examples of accounts of shared cooperative intention and JICA that embody the notional singularization move. In the course of presenting them, it will become clear why I think that they should be interpreted as accounts of shared cooperative intention and JICA, rather than of shared intention and joint intentional action more generally.

Note that accounts that *allow* the parties of a shared cooperative intention to each intend "to form and enact a single system of belief and desire" or to do something as some kind of group agent, but which also include conditions that are supposed to independently ensure intentional cooperation, do not exemplify the notional singularization move (e.g. List & Pettit, 2011, p. 34). After all, the parties' conception of themselves as parts of a single agent or body is supposed to make such conditions superfluous.

## 3.1 Jointly committing to intend as a single body/agent

According to Gilbert, several agents intentionally act together if they act on a joint commitment to intend as a single body to perform an action. A joint commitment, on Gilbert's view, is a commitment that has several agents as a subject. These agents together create the joint commitment by all expressing their readiness to jointly commit to emulate having a psychological state as a single body under conditions of common knowledge, such as, for example, to intend as a single body to *A* or to believe as a single body that *p*. Once the joint commitment has been created, the parties have certain non-moral directed obligations and entitlements vis-à-vis each other (Gilbert, 2008, pp. 506-508).

Gilbert describes the joint commitment to intend as a single body or, to put it differently, to espouse a goal as a single body as follows:

> The relevant joint commitment is an instruction to the parties to see to it that they act in such a way as to emulate as best they can a single body with the goal in question. The idea of a single goal-endorsing body is not itself understood in collective terms. The concept of a body that endorses a goal is neutral with respect to the nature of the goal endorser and with respect to its composition. (Gilbert, 2002a, p. 67)

The parties are not jointly committed to actually constituting a single body with the goal or intention to do something. Rather, they are jointly committed "*as far as possible to emulate, by virtue of the actions of each, with respect to its intending, a single body* that intends to do the thing in question." (Gilbert, 2008, p. 503) However, Gilbert uses the term 'body' loosely here: "One might substitute in the previous formulation the term 'person', for instance, or 'agent'." (ibid.) And: "One might use other phrases such as 'as a unit' or 'as one.'" (2002a, p. 67) In light of this loose usage, emulating could in some cases arguably amount to constituting.

Gilbert does not say why the content of the joint commitment includes the notion of intending as a single body (or person, agent, unit etc). But part of the motivation can be gleaned from her work on collective belief. According to Gilbert (1987; 1996), a collective belief is constituted by several individuals' joint commitment to believe something as a single body. She convincingly argues that an everyday claim that a group has a belief or view that *p* normally does not entail that the members of that group believe that *p* (e.g. the hiring committee may believe that Charlotte is the best candidate even if no committee member personally believes this). Now, the idea that the members see themselves as if they were parts of a larger body or distinct agent

helps make sense of a potentially radical discontinuity between the group and its members:

> [S]uppose someone claims that Jill believes that it is raining. The fact that Jill's toe does not believe it is raining or—to be fanciful—that Jill's toe believes it is not raining would not, surely, militate against the truth of this. Jill is one thing, her toe is another. Similarly, the single person or body we are committed together to constitute, as far as possible, is one thing. Each of us is another. (Gilbert, 2005, pp. 35–36)

According to Gilbert, there is a similar discontinuity between a group's shared cooperative intention and the members' personal intentions. This is why she also appeals to the notion of doing something as a single body or agent in her account of shared cooperative intention. When a joint commitment to emulate a single body that intends to *A* is in place, each party has an individual commitment to do his or her part of that emulation. However, this is not the personal commitment of an intention.[5] The individual commitment is derived from the joint commitment and depends on its existence. For the shared cooperative intention to appropriately result in a JICA, each party must furthermore act in light of this individual commitment when he or she does his or her parts (Gilbert, 2002a, p. 68).

---

[5] However, it is "likely" and "predictable" that the parties to a shared cooperative intention will form concordant personal intentions (Gilbert, 2008, pp. 509-510).

In Gilbert's discussions of shared cooperative intention and acting together, the adverbial 'as a body' primarily enters into the characterisation of the intending (the emulated psychological state), rather than into the characterisation of the acting (the content of the emulated intention). To illustrate, in discussing the case of Sally and Tim who jointly commit to intend as a body to go for a walk in the afternoon, Gilbert unpacks the content as follows: "Sally and Tim are jointly committed to intend as a body to produce, by virtue of the actions of each, *a single instance of going for a walk with the two of them as the participants in that walk*." (2008, p. 503) This way of construing the content of the emulated intention may make the account look viciously circular. After all, part of what an account of shared cooperative intention is supposed to help us understand is precisely what it is for several agents to be participants in one intentionally cooperative joint intentional action.

Gilbert can avoid circularity by appealing to a concept of a single instance of going for a walk with several participants that is, as Bratman puts it, ''*cooperatively neutral*'' (1992, p. 330). Indeed, one of Bratman's examples of such a concept would suit the case of Sally and Tim well: "We have, for example, a concept of our walking down the street that involves only the ideas that, roughly, we are each intentionally walking down the street, that our walking is alongside each other and at a comparable pace, and that we are each avoiding collisions with the other." (2014, p. 46) Another way of avoiding circularity would be to appeal to the notion of doing something as a single body also in characterising the content of the emulated intention. Indeed, for several parties to successfully continue to emulate a single body that is intending to go for a walk once the time of action has arrived, they must arguably emulate a single body *that goes for a walk*. In line with this, when Gilbert presents a case where Jane

and Hilda form a shared cooperative intention to hike to the top of a mountain, she writes: “They thus jointly commit to intend as a body to hike (as a body) to the top of the mountain.” (2006, p. 105)[6]

Let us return to Gilbert’s motivation for employing the adverbial ‘as a single body’ in her account. Besides blocking entailments from what the group intends to what members intend, Gilbert remarks that there are “more positive aspects of [the notion of] *X*-ing as a body.” (1996, p. 349) Plausibly, an implied cooperativeness between the parties of the joint commitment is supposedly one such positive aspect.

Such an implied cooperativeness is needed in order for Gilbert’s account to capture the intentionally cooperative character of JICA. The commitment to the emulation occurring “by virtue of the actions of each” does not itself rule out brute coercion. Suppose that Jane and Hilda hike to the top of the mountain by way of Jane’s grabbing, lifting and carrying Hilda against Hilda’s will as Jane herself walks forward and Hilda struggles in her arms. Hilda’s will is that they emulate the hiking of a single agent or body in a way that involves the actions of each. However, the way in which she wills that they do this differs from the way in which Jane wills that they do it.

---

[6] Gilbert (2007) also writes the following regarding five people who are acting on a joint commitment to intend as a single body to push a car up a hill: “An observer might say that they were behaving as if they constituted not five separate two-armed, two-legged individuals going about their personal business, but rather one ten-armed, ten-legged individual going about *its* business. This is how their joint commitment instructs them to appear, as far as is possible.” (p. 158)

Hilda's will is thwarted here but both may nevertheless be acting on their individual commitments to do their part of emulating a body that hikes to the top of the mountain. Even if their wills are not co-realisable, they can be bound together by the joint commitment.[7]

Neither is it clear that an appeal to "background social conventions" for how to carry out a joint activity will rule out brute coercion (Gilbert, 2008, p. 504). While the way in which Jane and Hilda hike "together" in my preceding description does not conform to the conventional form that hiking together typically takes in contemporary Western culture, we can imagine a (fanciful) brute coercion-involving case that does so conform. Switch to Sally and Tim: Suppose that they go for a walk together, each intentionally walking alongside the other down the street at a comparable pace and avoiding collisions with the other. Furthermore, Sally is pulling strings attached to Tim's legs, ensuring through brute coercion that they walk along her preferred particular path (this is compatible with Tim's walking being intentional under a coarse-grained description). Or suppose that Tim is strapped into a powered exoskeleton that is under Sally's control and that Sally is partly controlling the way they walk by brutely coercing some of Tim's bodily movements. It would not do to say that the conventional form of walking together is to do so as a JICA or in an intentionally cooperative way, since that would make the account viciously circular.

---

[7] *Too* brute coercion would be ruled out though. Hilda must in some way be able to act on her joint commitment-derived individual commitment, so Jane could not knock Hilda unconscious and just carry her limp body up the mountain (cf. Hayek 1960/2011, pp. 199-201).

If anything is supposed to rule out brute coercion in Gilbert's account of shared cooperative intention and acting together, it is the notion of intending as a single body. Admittedly, Gilbert never explicitly says that her account is meant to rule out brute coercion and she never discusses any brute coercion-involving cases. But in comparing her and Bratman's accounts, she argues that a shared cooperative intention on her account provides "a more felicitous, because more stable framework for bargaining and negotiation and, relatedly, a more felicitous means of coordinating the personal intentions of individuals, and keeping them on the track of the shared intention." (ibid., p. 510) It is more stable and felicitous because, unlike personal commitments, joint commitment-derived individual commitments are not unilaterally rescindable. A joint commitment can only be rescinded by the parties together. If a party fails to fulfil his or her individual commitment, then the others have a right to rebuke him or her. Hence, her account is supposed to be able to play the same functional role as Bratman's interlocking intentions and intended meshing of subplans—that is, as those elements in Bratman's account in virtue of which it succeeds in ruling out intentional brute coercion. Furthermore, the absence of any discussion of cases of brute coercion in her various exchanges with Bratman suggests that she agrees with him that such cases should be ruled out.

## 3.2 Identifying as an agent's part—reasoning as the whole

The notional singularization move is also exemplified in team reasoning-based accounts of JICA. An individual engages in team-directed reasoning when, in a context involving other agents, she asks herself 'What should we do?' rather than

'What should I do?'. What I am interested in are attempts to provide accounts of shared cooperative intention and JICA based on team-directed reasoning that is prompted in the participants by group identification and framing (Bacharach, 2006, pp. 137–41; Pacherie, 2011; 2013). What I will highlight is the conception of the group or team—the 'we'—that is supposed to be involved in this group identification and framing.

Pacherie's (2013) account is the most explicitly articulated available team-directed reasoning-based account of shared cooperative intention. On this account, several agents' intentions to do their parts to bring about a single "group goal" *A* constitute a shared cooperative intention to *A* if they are each formed as a result of team-directed reasoning.[8] What each participant asks him- or herself is "Which single group goal, $A_1$ or $A_2$ or … $A_n$, should we bring about?". The selected group goal is thus the outcome that the shared cooperative intention is directed toward. Hence, the shared cooperative intention is downstream of team reasoning.[9]

---

[8] The group goal is distinct from the "team's objective" (Bacharach, 1999; Sugden, 2003; Gold and Sugden, 2007). The team's objective is that relative to which the goodness of available outcomes is ranked and to which the team utility function is defined. A football team's objective could be to score a point against the opposing team, whereas the group goal of two team-mates at a particular moment might be that they make a pass play to the left.

[9] A shared cooperative intention could also be upstream of team reasoning, perhaps settling the team's objective and specifying the team utility function (Gold & Sugden, 2007, pp. 117, 136; Hakli et al., 2010).

Following Bacharach, Pacherie takes team-directed reasoning to be primarily the result of arational group identification and the framing of a decision problem as a problem facing the group or team.[10] A team-directed reasoner frames a coordination problem as a problem for him- or herself and the other agents considered as parts of a single agent. According to Bacharach, a decision-maker who identifies with a group or team thinks of him- or herself as "part of some collective doer" (2006, p. 137). Similarly, Pacherie (2013) takes a group-identifying "player" to "conceive of this group [that he or she identifies with] as a unit of agency acting in pursuit of some group goal." (p. 1832) This is in line with how proponents of team reasoning theories in general think about the group or the team. The 'we' in the team-directed reasoner's question should be understood in a particular way, namely as "a unit of agency, acting as a single entity in pursuit of some single objective" (Gold & Sugden, 2007, p. 125).

On this conception then, identifying with the group or team is identifying with *an additional agent* over and above the group members themselves. Indeed, Bacharach sometimes expresses himself in a way that suggests that each player sees him- or herself and the others as individuals who subordinate their choices to an "imaginary manager" or "director" who gives instructions to each regarding what to do (1999, p. 118; 2006, p. 123).

Why think that we ever deliberate and act in this way? The reason is that the question 'What should *I* do?' yields no determinate rational answer in some simple

[10] This raises difficult questions about if and how team-directed reasoning can be rationally justified (see Roth, 2014).

coordination games that intuitively do have unique rational solutions. Consider the Hi-Lo game, here narrated as The Footballers' Problem (Sugden, 2003): Two players on the football team are trying to make a pass play. In the heat of the game, they cannot communicate. The pass play can be made to the left or to the right of the receiving player. A pass to the left would be preferable. A pass play to the left will be brought about if player 1 passes the ball to the left while player 2 runs to the left to receive it. Least preferable is a failure of coordination. The "game" can be represented as follows, where numbers represent utility for each player:

| | Left | Right |
|---|---|---|
| Left | 10, 10 | 0, 0 |
| Right | 0, 0 | 5, 5 |

Orthodox game theory provides no determinate rational solution to this game. Each player is supposed to choose the best response to whatever she believes the other player will do. The theory tells each player: *if* the other plays Left, then play Left; if the other plays Right, then play Right. But whether the other plays Left or Right depends on what the other thinks that the player him- or herself will do. There is no factor that can rationally tip the players' expectations about whether the other will go Left or Right. But, for each player, passing/running Left is intuitively the obviously rational thing to do! Indeed, without viewing the situation through the theoretical lens of game theory, it is hard to see that there is a problem at all here.

Team-reasoning football players overcome this by each first selecting the outcome that is best for the group agent that they take themselves and the other to be parts of. Each player evaluates the outcomes in light of her team-directed preferences rather than her private preferences. It is typically assumed that the participants share a single shared team utility function that is common knowledge between them, and that the team utility is equal to the average of their expected private utilities. The team reasoning Footballers' Problem could thus be represented like this:

| | Left | Right |
|---|---|---|
| Left | 10 | 0 |
| Right | 0 | 5 |

Each then intends to do his or her own part of the action profile—in this case pass/run Left—that is likely to bring about the outcome that is best for the group agent. The intention of each is their "participatory intention" (Gold & Sugden 2007, pp. 111, 121, 126).[11] For example, one of the football player's participatory intention will be to pass the ball to the left, which is his part of the action profile that also includes the other player's action of running to the left. A theory of team reasoning can thus vindicate our intuitive judgement that each player should pass/run Left.

[11] Gold and Sugden prefer 'collective intention'. Tuomela (2007) and Ludwig (2016) prefer 'we-intention'.

This solution is in one sense entirely individualistic. Team reasoning is nothing more than processes of *team-directed* reasoning that each player could engage in unilaterally rather than in parallel. Similarly, team preferences are nothing more than a set of shared *team-directed* preferences according to which each player could unilaterally evaluate outcomes (Sugden, 2000).[12]

In making sense of group-identification and framing, Bacharach and Pacherie draw on socio-psychological research on social identity and group membership, as well as research on the perception and categorisation of groups by out-group members. In the literature on the perception and conceptualisation of groups, one indeed finds the idea that people sometimes think about and respond to groups as if the groups were larger single agents. In an influential paper, Robert Abelson, Nilanjana Dasgupta, Jaihyun Park and Mahzarin Banaji (1998, p. 248) ask which cues make people see a group "as capable and motivated to act as a purposeful unit." According to Marilynn Brewer (2015), people often think about groups in terms of an implicit "dynamic agency theory". When a group is conceptualised with this theory, "attributions [of motives or goals for example] are made to the group as a collective actor […]." (ibid., p. 164) Thinking about a group as a unit in this way "leads to responding to the group as if it were a single individual, but on a larger scale." (ibid., p. 166)

Bacharach (2006, pp. 70-73) takes the conception of the group that an agent has as a result of group-identification to be the same as that which she would have of it as an

[12] Note that team-directed reasoning is *not* what Bacharach (1999) calls "group benefactor reasoning".

external out-group perceiver. In both his and Pacherie's view, group-identification involves thinking and reasoning about the group that one identifies with as a single entity—in other words, it involves "entifying" the group from the inside. Entification from the inside is thus just entification from the outside plus the accompanied transformation of the group-identifying agent's self-conception.[13]

Aside from this influence from social psychology, the more principled reason why proponents of team reasoning theory operate with their peculiar conception of a group or team as a single agent is arguably the thought that this conception brings with it a sense of cooperation.[14] This would then be an instance of the agent version of the notional singularization move. Pacherie's (2013) account targets what she calls "strongly jointly intentional" joint action, in which participants "should think of [the joint action] as a collaborative activity." (p. 1822) She takes group identification and the accompanying framing of a decision problem in we-terms to provide a cognitively undemanding way of imbuing participatory intentions with a cooperative character (2013, p. 1834; 2011, pp. 178, 187).

---

[13] This conception of the group as a single agent is absent from much of the literature on social identity. It cannot be found in Hogg and Abrams' (1998) advanced textbook on social identity theory for example.

[14] One might think that there is another principled reason, namely that "an individual agent can only select her strategy, but a group agent can in a sense select outcomes" (Hakli et al., 2010, p. 298). However, it is not clear what to make of this within the framework of Bacharach's and Pacherie's accounts since no actual group agent is posited in their accounts: there are only individual agents, each of whom engages in team-directed reasoning.

## 4 Notional agent singularization and why it fails

The argument against the agent version of the notional singularization move is simple: Nothing rules out that the goal-directed activity of the single group-level agent, of which each takes him- or herself to be part, is implemented through brute coercion and conflict between those parts. For example, think of the activity among the employees of a corporation, the citizens of a nation state, or the activity of insects that together make up a "swarm agent". Suppose the dictator of a totalitarian state lets cabinet officials fight each other to death in order to determine what the state's official policy on some issue should be for example.[15] Or, to take a more benign example, the actions of a corporation may be partly implemented by some employees' brutely controlling the behaviour of other employees through the use of security guards, the locking of doors and physical restriction of access to certain information, and so on. What matters for the agency of the whole are the global properties of the behaviour of the collective. There is no need for this behaviour to be implemented by cooperative interactions between the parts. What matters is simply that each part is playing the right causal role or fulfilling the right function within the system. Similarly, an additional agent, such as an "imaginary manager" or "director" (Bacharach, 1999, p. 118; 2006, p. 123), could be imagined to instruct individuals to brutely coerce each other to bring about what is best for him- or herself.

---

[15] Arguably, instituting a dictatorship is a way to collectivize reason in a group so that it can be appropriately treated as an intentional agent of its own (List & Pettit, 2011, p. 53).

My claim is not, of course, that two people who make a pass play, carry a table or go for a walk together are anything like a corporation or a totalitarian nation state. On the contrary, the analogy between the activity of a group-level agent such as a corporation or a national state and JICA is misleading. The point is rather that it is logically and conceptually possible for there to be a single higher-level agent whose intentional actions are implemented by coercive interactions between parts of the supervenience base of these intentional actions. Corporations, nation states or swarms of insects are merely meant to illustrate this possibility. The possibility, in turn, shows that agents that reason team-directedly or that are jointly committed can think of themselves as parts of a higher-level agent and rationally act to bring about the goal of that higher-level agent, while at the same time trying to coerce the other participating agents into doing so.

One might think that the problem with the notional agent singularization move is that the parties of shared cooperative intentions think of themselves as if they were parts of the supervenience base of a *group-level* intentional agent. What leaves room for brutely coercive intentions between these parts, one might think, is the introduction of different levels of intentional agency: the level of the imagined group agent and that of the participants. However, we will see that relying on at least one much weaker notion of a single group agent does not help.

According to Björn Petersson, getting "the distinction between a shared intention and a set of individual noncooperative intentions" right, and providing conditions for shared intention such that they "suffice to ground [a sense of] collectivity" are one and the same task (2007, p. 141). He argues that we get the distinction right and ground the requisite sense of collectivity by requiring that participants of a joint

action think of themselves as parts of a single *causal group agent* of which the other agents are also parts. All that thinking of "we" as a causal agent involves is an implicit assumption that something "glues" the group members together. This is an assumed base of behavioural dispositions and causal powers attributed to the group or the "we". This assumption can "be expressed in terms of 'causal cooperation.'" *(ibid.,* p. 156)

Note that this "causal cooperation" is compatible with brute coercion. Consider Petersson's example, adapted from Davidson's, of one person who jiggles another's hand so that coffee is spilled from the cup that the other is holding. This illustrates a form of causal agency that is neither merely behavioural nor intentionality-involving:

> If you point out that your jiggling made *me* spill coffee, my causal role for that effect is stressed—although no intentions are in play. My spilling coffee is in this case something that I do, albeit unintentionally, so it is an act, but not in a stronger sense than the one outlined. (*ibid.*, p. 149)

A third person who observes the described incident could similarly stress the causal role of the group consisting of "you" and "I". This observer would point out that *they* spilled coffee. She would conceive of the two other individuals as parts of one causal agent, even if one of them acted on a coercive intention to bypass the other's intentions/subplans or treated the other as a mere tool or object. This conception would be appropriate, since nothing in Petersson's conception of a causal agent rules out that the brutely coercive intention constitutes the glue that makes *them* into a causal agent of the spilling of the coffee.

As I mentioned in section 2, singular composite intentional action is in a sense cooperatively coordinated at a time and over time. Could the intentionally cooperative character of JICA be accounted for by appeal to the intentionally cooperative character of the imagined larger agent's composite intentional action? No, this would be to succumb to the fallacy of decomposition—the fallacy of inferring that parts have a certain property *P* just because the whole has the property *P*. One cannot infer that the interaction between the parts of the group agent has an intentionally cooperative character just because the activity of the whole group agent has such an intentionally cooperative character over time.

I am assuming here that participants are supposed to think of themselves as if they were *spatial* parts of a single group-level agent. But perhaps the participants could instead think of themselves as if they were *temporal* parts of a single agent. In order to determine whether this could save the notional agent singularization move, we would need an account of singular composite intentional action. However, existing accounts of rational coordination and "cooperation" in a single agent over time would not be of help here. This is because they are themselves based on an analogy with intentional coordination and cooperation between multiple agents (e.g. Bacharach 2006, pp. 88-89, 191-198; Gold 2013). Hence, insofar as the point of the notional singularization move is to enable a relatively simple and conceptually undemanding account of JICA, it is not clear how the temporal parthood interpretation would help.

## 4.1 Gilbert's joint commitment-based account

Consider first Gilbert's joint commitment-based account in light of her well-known example of walking together. Given the notional agent singularization move, Jane and Hilda jointly commit, by virtue of the actions of each, to emulate a single agent that intends to produce a single instance of hiking to the top of the mountain in which the two of them participate. What about the notion of a single agent could help Gilbert rule out that Jane and Hilda hike together by way of, for example, Jane's grabbing, lifting and carrying Hilda against Hilda's will as she herself walks forward and Hilda struggles in her arms?

Arguably, nothing would. Their walking "together" in this way would be a rather accurate emulation of a single agent that intends to produce a single instance of hiking that involves both Jane and Hilda. It would only be inconsistent with Gilbert's account if the joint commitment were required to specify the content of the hiking movements of the single agent in very fine-grained detail, perhaps all the way down to the level of the basic actions of the parts played by Jane and Hilda (on basic actions, see Searle, 1980, pp. 65–66). However, it isn't clear why the parties would need to have such fine-grained joint commitments, and there is no reason to think that the typical communicative exchanges with which we establish joint commitments specify their content down to such fine-grained detail: Jane and Hilda might form their shared cooperative intention by Jane's asking Hilda, "Shall we hike to the top of the mountain?", and Hilda's responding, "Sure!" (Gilbert, 2008, pp. 487, 504). Given that Gilbert's account is interpreted as an instance of the notional agent singularization move, the implicit content of this exchange would not rule out brute coercion.

Gilbert could perhaps respond that some cases of brute coercion aren't supposed to be excluded by her account: It is not an account of shared cooperative intention and JICA, but of shared intention and joint intentional action as such. If Gilbert's account is an account of former though, it fails to provide sufficient conditions for its target phenomenon, at least given the notional agent singularization interpretation.

## 4.2 Group identification-based accounts

Group-identifying team reasoners coordinate their actions to reap the greater mutual benefit of the Hi equilibrium in Hi-Lo. Team-directed reasoning prompted by group identification may therefore seem like a process that leads to cooperative intentions. Indeed, given certain simplifying assumptions about the circumstances under which several individuals team reason, team-directed reasoning cannot lead to rational brute coercion between participants. In a sense, such coercion is therefore excluded. But as we will see, it is excluded because of assumptions that are external to the minds of the participants. It is not that the team reasoners are committed to keeping these assumptions true or that group-identification presupposes them to be true, at least not if the participants who identify with the group think of the group as a single agent.

Now, Bacharach assumes that to identify with a group is in part to presuppose the existence of a team objective and team utility function that is shared by all participants. This, at least, is implied by his working definition of a group (Bacharach, 2006, p. 87). He also assumes that "players" share a single set of team-directed preferences that is common knowledge between them (*ibid.*, pp. 123, 151). Given that

the target phenomenon here is shared cooperative intention, I think that Bacharach is right to take the participants to presuppose this (see the discussion of the notional body singularization move in section 5).[16] However, given Bacharach's view of group-identification, where the players conceive of the group as an imagined additional group-level agent, and see themselves as parts of that agent, it is not clear why the participants would presuppose this.

Bacharach also frequently assumes that team utility is equal to the average of the agents' private expected utilities, even though private and team-directed preferences are logically independent. If these assumptions are in place and if there is a unique rational solution to a coordination problem, then players who identify as parts of the same single group agent and engage in team-directed reasoning cannot rationally select distinct group goals or action profiles. Hence, in such circumstances the resulting participatory intentions will be complementary in such a way that brute coercion will be excluded.[17] Any intended brute coercion would have to be

[16] With the exception, perhaps, of the presupposed common knowledge.

[17] At least if the participants reason team-directedly all the way to an action profile that specifies participants' intentionally basic actions. Gold and Sugden (2007) seem to assume that they do (just as Bratman takes the intended meshing of subplans to reach all the way to action). Their account is supposed to provide an explanation of how the participants can rationally pursue intentions to make their "low-level, *tactical* intentions" co-realizable and achieve "mutual responsiveness in action" in some situations (ibid., p. 136).

compatible with the intention of each participant.[18] Hence, each participant's team-directed reasoning could not recommend an action profile that could be rationally executed in a way that involves intentional brute coercion between participants.

Here, we need to step back and consider what the relations are between theories of team reasoning, group identification, and accounts of shared cooperative intention. I take it that the relations are the following: First, the fact that team reasoners would be able to solve the Hi-Lo problem in the context of game theory, combined with the fact that we are actually able to rationally solve Hi-Lo-like problems in real life, suggests by inference to the best explanation that we are indeed team reasoners. Secondly, it is plausible that group-identification will trigger individuals to engage in team-directed reasoning. If this is right, then we are free to appeal to this team-mode of practical reasoning also in contexts where some assumptions standardly made in game theory do not hold.

In real life, the assumptions often do not hold. For example, the game may not be common knowledge between the participants, or the players may have different team-utility functions. In addition, it is unclear to what extent there is a link between the socio-psychological phenomenon of identifying with a group and thinking that every member of the group has the same team utility function or that all agree on what the team's objective is (for an overview of the socio-psychological literature, see Hogg

[18] This does not rule out that non-brute coercion can be involved in the emergence of a team utility function, potentially resulting in a function that does not fairly reflect some team members' private utilities.

and Abrams, 1998). At least, there is no such conceptual link if the group is conceived of as a large single agent and the members as its parts. Furthermore, the dependences between private and team-directed utility can potentially be very complex; higher team-directed utility need not reflect higher private utility (Sugden, 2000, pp. 178–79; Tuomela, 2007, pp. 159-163; Pacherie, 2013, p. 1836).

If these various simplifying assumptions are relaxed, rational team members could knowingly act to bring about distinct group goals or actions (Hurley, 2005, pp. 593–94). Furthermore, if group goals or actions are distinct, then one player may use brute coercion to make sure that the other's contributory action is performed. Suppose that you and I each intend to do our respective part in bringing about the group goal of moving aside a heavy table (Pacherie, 2013, p. 1836). For each of us, the group goal or action is the output of our own team-directed preferences and team-directed reasoning. Furthermore, suppose that my team-directed preferences and reasoning dictate that the following outcome is best for the imagined single unit of agency of which I take us to be parts: that we walk one in front of the other, carrying the table between us with the long sides of the table aligned with the walking direction. Your team-directed preferences and reasoning, on the other hand, dictate that the following outcome is best for the single unit of agency of which you take us to be parts: that we walk alongside each other, with the long sides of the table perpendicular to the direction in which we walk. In this sort of situation, there is nothing akin to the required interlocking of intentions and subplans that would make it irrational for either of us to push, obstruct or otherwise coerce the other by brute physical force into doing a part of bringing about the single group agent's goal, that is, the other part of the group action.

On the assumption that the participants are rational, this kind of case can be ruled out if the right kind of relationship between team-directed and private preferences is assumed or stipulated. Outcomes that may have to be brought about by coercion will generally have lower expected private utility than outcomes that will be brought about non-coercively. This would, one might think, be reflected in lower expected team-directed utility. However, given that Bacharach, Pacherie, and Gold and Sugden all make use of the notion of the group as an imagined additional group-level agent, there is no reason to assume a relationship between team-directed and private preferences which ensures this, nor is there any reason to think that brute coercion would be excluded.

To conclude, neither the team-directed mode of practical reasoning nor the group-identification and framing that trigger it result in intentions with a cooperative character. What rules out brute coercion are rather the simplifying idealised assumptions that team reasoning theorists and game theorists take for granted. These are primarily assumptions made by the observing theorist. Only derivatively are they assumptions of the participants, as a result of their having common knowledge of the game. Both instances of the notional agent singularization move that I have reviewed thus fail.

In the next section, I argue that something like the simplifying idealised assumptions of team reasoning theorists are nevertheless important. They help explain why a suitably qualified form of a notional body singularization move can be made to work. However, the assumptions are then assumptions of the participants, assumptions that frame their participatory intentions.

# 5 Notional body singularization and why it works

The body version of the notional singularization move is intuitively more promising than the agent version. A living body in action is normally such that the body parts coordinate and mutually adjust to achieve a task or goal in an efficient and comfortable way. It is an integrated collection of components that in some sense cooperate to reach some end state.

For the notional body singularization move to work beyond a narrow and artificially restricted range of cases of JICA, we must not take the notion of a single body too literally (see also Butterfill, 2017, p. 476). In short-form improv comedy there is a "game" where the improv ensemble members all try to simultaneously talk about a suggested topic as if they had one common mouth and voice. But such cases, where participants try so far as possible to intend or act as if they literally were a single body, are rare. On the other hand, as we have seen, the notion of a single body cannot just be vaguely equated with the notion of a single agent or unit.

We must also appeal to something more specific than a single body in action. This is because an action or activity performed by a single body can involve something like brute coercion. Consider, for example, a speleologist who has to go through a very tight passage. Suppose that the most efficient (and painful) way of doing this is to use the right arm to force the left arm and shoulder through the narrow opening. Here, this involves parts of the body "coercing" other parts for the sake of a selection of optimal

means given a certain situation.[19] A single body can also in a more permanent manner use a kind of coercion in order to successfully achieve coordinated action. Suppose that I suffer from spasms in parts of my body. I have learned to control these spasms by exercising a kind of brute force over the rest of my body, so that I successfully manage to walk down the street despite the spasms.[20] It seems that, in both these cases, the action is that of the whole single body, rather than that of the whole-body-minus-the-coerced-parts. In light of this, we thus need to appeal to a *properly functioning single body in action* rather than to a single body in action *tout court*.

One might suspect that this proposal just reflects a myopic anthropocentrism. Could a single body in action not be such that its different parts were competing with or coercing each other when it functioned properly?[21] Indeed, once we look at non-human bodies around us, one might think that there are such bodies. Consider Jonathan Dale's (1999) model of how the coordinated movement of a starfish is produced. The movement is the result of a mutual inhibition of motor centres in the rays of the starfish (its "arms"), which compete for dominance. There is no centralised control according to the model. However, this does not show that there is no conceptual connection between a properly functioning body in action and brute coercion-excluding coordination between the body's parts. The fact that a competitive dominance algorithm is implemented in the nervous system of the starfish does not

---

[19] I am grateful to an anonymous reviewer for this case.

[20] Thanks to Michael Bratman for challenging me with this case.

[21] Thanks to an anonymous reviewer for raising this objection.

mean that its rays—the parts of the controlled system that are being coordinated—are competing for dominance or that they are involved in mutual coercion.

My suggestion is that there is a canonical concept of a properly functioning body in action that does imply efficient coordination between body parts. Perhaps organisms can be found whose biological boundaries do not coincide with a single body in action according to such a canonical concept, but this need not be a problem for my proposal. As long as the human participants are themselves myopic anthropocentrists and possess the canonical concept, the notional body singularization move can be made to work. It is this concept that is relevant for specifying the required content of the participants' commitments or intentions.

The notion of a properly functioning single body in action must not tacitly involve the very notion of cooperation that we are trying to illuminate. Such a vicious circularity can be avoided if we understand cooperative coordination of body parts in terms of the most rational or efficient bodily movement with respect to a goal, given the situational constraints. Recall from the previous section (4.2) how the simplifying assumptions made by team reasoning theorists make rational brute coercion impossible: Given a single commonly known team utility function—defined relative to a single team's objective—according to which team utility equals the average of the players' private utilities, the implementation of an action profile that involves brute coercion will from the team-directed reasoner's point of view always be less preferable than an available alternative implementation that doesn't. If these assumptions hold and if the rational choice is the team utility-maximising choice, then a player could not appropriately choose an implementation that involves brute coercion. Because of the single team utility function, the players' actions will be

complementary in such a way that brute coercion will be excluded. Any brute coercion would have to be compatible with the actions of the others. My proposal is that, unlike the notion of a single agent in action, the notion of a properly functioning single body in action implicitly involves something like these assumptions, where the "players" are parts of the bodily whole rather than autonomous intentional agents.

The assumptions implicitly involved in the canonical concept of a properly functioning body in action allow us to interpret movements of another bodily agent as making up a purposive action directed at a specific goal, and to interpret its behaviour as involving the selection of means to realize that goal. According to Gergely Csibra and György Gergeley (2007), assumptions like these often underlie our interpretation of a body in action. In particular, they argue that young children understand the behaviour of other bodies according to a teleological interpretative scheme based on an "efficiency principle" (ibid., p. 70). This interpretative scheme "presupposes that (1) actions function to bring about future goal states, and (2) goal states are realized by the most rational action available to the actor within the constraints of the situation." (Gergeley & Csibra, 2003, p. 289) The scheme is teleological rather than mentalistic; it does not involve representations of mental states. The "goal state" here refers to the outcome that the behaviour's function is to bring about, and the situation that provides the constraints is the actual situation rather than what the actor believes the situation to be. Prima facie, it is plausible that adults and not only young children often interpret the behaviour of others using this teleological interpretative scheme, even if adults in many situations can also interpret a mentalistic scheme that explains the actor's behaviour in terms of ascribed mental states.

The most rational course of action for a bodily agent is the pursuit of the most efficient means (with the highest utility) of realizing the agent's goal. When it comes to bodily actions of human beings, this will be the least energetically costly profile of the movements of body parts and segments needed to realize the bodily agent's goal given the constraints of the situation. We assume that different body parts are not working at cross-purposes in interpreting the behaviour. Arguably, we presuppose that there is a single shared utility function for all coordinated parts. Now, given that the efficiency principle is implicit in our interpretation of an individual properly functioning single body in action, this principle could also be applied in the interpretation of a larger single body in action (cf. Mascaro & Csibra, 2014; Török, Pomiechowska, Csibra & Sebanz, 2019).[22]

According to a body version of the notional singularization move, these implicit assumptions feature in the content of the participants' commitments (Gilbert) or in their framing of themselves in the context of a decision situation (Bacharach, Pacherie, Gold and Sugden). The concept of acting as part of a properly functioning single body in action enables a kind of minimally cooperative intent. Acting on this kind of intent rules out that participants intentionally and brutely coerce co-participants, that is, the participants cannot knowingly treat the others merely as tools

[22] Olivier Mascaro and Gergely Csibra (2014) report the preliminary results of an experiment meant to test whether the efficiency principle is also at work in children's understanding of joint goal-directed action. The results suggested that young children interpreted the behaviour of two creatures on a screen on the assumption that the creatures were together choosing the collectively most efficient joint means of realizing a shared goal.

or objects to be manipulated. It does this because the notion of a properly functioning single body in action carries with it an implicit assumption that there is something like a single shared utility function of all body parts. The coordination of the parts is thus understood with respect to a single task or goal of the whole (corresponding to the team-reasoning theorist's "team's objective") relative to which the shared utility function is defined.

Talk of a shared utility function may seem misplaced in the context of Gilbert's joint commitment-based account of shared cooperative intention. Nevertheless, I think it can illuminate how the concept of a properly functioning single body in action implies a kind of minimal cooperativeness. Consider again Jane and Hilda's joint commitment to intend as a single body to go for a walk. Each is committed to as far as possible emulate that they intend—as if they were parts of a properly functioning single body in action—to go for a walk that involves both of them as participants. With such content, Jane and Hilda's joint commitment would not be satisfied if Jane tried to implement the commitment by way of her grabbing, lifting and carrying Hilda against Hilda's will as she herself walks forward. Suppose that this action is part of an action profile that is associated with an outcome with the highest expected team utility from Jane's point of view. If the team utility function according to which this outcome has the highest expected utility is assumed to be shared, then Hilda will not resist Jane but rather let herself be carried forward. At least, she will do so if she is rational.

Crucially, the joint commitment that they intend as a properly functioning body to produce a single instance of walking in which they both take part will not be satisfied if their team utility functions differ in such a way that one participant ends up brutely

coercing the other. The way in which the goal is achieved must maximise a utility function that is shared between the participants, otherwise their shared cooperative intention—the joint commitment to intend as a properly functioning single body in action—will not be satisfied. Furthermore, the participants are jointly committed to making it so that they have a shared team utility function, since this is part of what is implicit in the notion of doing something as a properly functioning single body in action.

Things are, perhaps surprisingly, less straightforward when it comes to the team reasoning-based accounts. In these accounts, the participatory intention of each participant is arguably simply the intention to do his or her own part of the action profile. Each participant thus does not *intend* that they espouse the team's objective as a single body in action (nor, of course, do they together jointly commit to espousing it as in Gilbert's account). Intentions to brutely coerce other participants are thus not ruled out by virtue of being inconsistent with one's participatory intention. Rather, each *conceives of* him- or herself and the other participants as parts of a properly functioning single body in action and judges that an action profile is such that it maximises expected team utility. This conception is part of the basis on which each participant reasons team-directedly and forms the participatory intention; it "frames" or "cognitively guides" his or her reasoning and the resulting intention (see Alonso, 2017).

If the conception exclusively has a mind-to-world direction of fit—if it merely consists of a belief or belief-like state—then it is nevertheless conceivable that a

rational participant could satisfy his or her participatory intention by brutely coercing another participant.[23] Consider the example from section 2 in which we are cooking together, the oven door breaks open at a crucial moment and, in accordance with my participatory intention, I quickly lift you up and shove you against the oven door to keep the heat from escaping. In such a situation, even if you try to resist, and even if it is expected that you will try to do so, this may be the most efficient and rational option given my team-directed utility function. Suppose I then discovered that you didn't act in light of this supposedly shared team-directed utility function. In this case, no intention of mine needs to have been frustrated, and my participatory intention may be fully satisfied, even if I brutely coerced you as I was satisfying it.

However, the conception on the basis of which I engaged in team-directed reasoning and formed my participatory intention turned out to have been bad or mistaken. There is a kind of error or mistake here, even if it is not one that makes me guilty of irrationality. In light of this, it is appropriate to characterise my participatory intention as having a cooperative character, given that it is framed by my (mistaken) conception of you and me as parts of a properly functioning single body in action, where this implicitly implies that there is a shared team utility function on the basis of which we act. Since the intention is formed by a process of reasoning about what this properly functioning single body in action should do, my participatory intention couldn't be intentionally brutely coercive (even if it could be satisfied by an action that happened to involve brute coercion).

---

[23] As Bratman observes, "one might, after all, *conceive* of G as a 'unit of agency' while wanting and intending that G not function in this way." (2014, p. 182 n. 19)

## 6 Conclusion

This notional singularization move fails in its agent version. There is nothing essentially cooperative about each participant's treating him- or herself and the others as if they were parts of a larger single agent. Each can do this while trying to brutely coerce the others in way that is incompatible with JICA. However, if the notional singularization move is interpreted differently, so that participants think of themselves and the others as if they were parts of a properly functioning single body in action and commit to realising the goal of a joint action as this single body, then the intentionally cooperative character of JICA can be captured. This is because the notion of a properly functioning single body in action implicitly includes both the assumption that the body's parts will be coordinated in the most efficient way to achieve a goal and the assumption that all those parts have a shared utility function. An account of shared cooperative intention that does not rely on agents' having interlocking higher-order intentions is thus possible.

## Acknowledgements

This paper has gone through many rewrites. I have received helpful questions, comments and suggestions from many people along the way, including Facundo Alonso, Adrian Alsmith, Franz Altner, Line Edslev Andersen, Jonathan Birch, Michael Bratman, Stephen Butterfill, Hein Duijf, Margaret Gilbert, Natalie Gold, Raul Hakli, Bryce Huebner, Holly Lawford-Smith, John Michael, Michiru Nagatsu, Carlos Nunez, Björn Petersson, Carlo Proietti, Alessandro Salice, Glenda Satne, Anne

Schwenkenbecher, Thomas Szanto and several anonymous reviewers for *Synthese*, *Ergo* and *Mind & Language*. I especially want to thank Margaret Gilbert for walking together with me in Vienna whilst discussing coercion and acting together (thanks also to Herlinde Pauer-Studer for inviting me there). I am also grateful to audiences in Copenhagen, Cork, Helsinki, Linköping, London, Lund, Kraków, Milan, Santiago and Warwick. The research was funded by project grant 421–2014–1025 from Vetenskapsrådet (the Swedish Research Council).

## References

Abelson, R. P., Dasgupta, N., Park, J. & Banaji, M. R. (1998). Perceptions of the Collective Other. *Personality and Social Psychology Review*, *2*(4), 243–250.

Alonso, F. (2017). Intending, Settling, and Relying. In D. Shoemaker (Ed.), *Oxford Studies in Agency and Responsibility: Volume 4* (pp. 50–74). Oxford, UK: Oxford University Press.

Bacharach, M. (1999). Interactive Team Reasoning: A Contribution to the Theory of Co-operation. *Research in Economics, 53*(2), 17–47.

Bacharach, M. (2006). *Beyond Individual Choice: Teams and Frames in Game Theory*. Princeton, NJ: Princeton University Press.

Bayles, M. (1972). A Concept of Coercion. In J. R. Pennock & J. W. Chapman (Eds.), *Nomos XIV: Coercion* (pp. 16-29). Chicago, IL: Aldine-Atherton.

Blomberg, O. (2016). Shared Intention and the Doxastic Single End Condition. *Philosophical Studies*, *173*(2), 351–372.

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

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

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

Brewer, M. B. (2015). Motivated Entitativity. In S. J. Stoessner & J. W. Sherman (Eds.), *Social perception from individuals to groups* (pp. 161-176). New York, NY: Psychology Press.

Butterfill, S. (2017). Review of Joint Commitment: How We Make the Social World, written by M. Gilbert. *Journal of Moral Philosophy*, *14*(4), 475–478.

Csibra, G. & Gergely, G. (2007). 'Obsessed with Goals': Functions and Mechanisms of Teleological Interpretation of Actions in Humans. *Acta Psychologica*, *124*(1), 60–78.

Dale, J. (1999). Coordination of Chemosensory Orientation in the Starfish *asterias forbesi*. *Marine and Freshwater Behaviour and Physiology*, *32*(1), 57-71.

Gergely, G. & Csibra, G. (2003). Teleological Reasoning in Infancy: The Naive Theory of Rational Action. *Trends in Cognitive Sciences, 7*(7), 287–292.

Gilbert, M. (1987). Modelling Collective Belief. *Synthese*, *73*(1), 185–204.

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

Gilbert, M. (1996). More on Collective Belief. In *Living Together: Rationality, Sociality, and Obligation* (pp. 339–360). London, UK: Rowman & Littlefield.

Gilbert, M. (2002a). Acting Together. In G. Meggle (Ed.), *Social Facts and Collective Intentionality* (pp. 53–71). Frankfurt, Germany: Hansel-Hohenhausen.

Gilbert, M. (2005). Shared Values, Social Unity, and Liberty. *Public Affairs Quarterly, 19*(1), 25–49.

Gilbert, M. (2006). Who's to Blame? Collective Moral Responsibility and Its Implications for Group Members. *Midwest Studies in Philosophy*, *30*(1), 94–114.

Gilbert, M. (2007). Acting Together, Joint Commitment, and Obligation. In N. Psarros & K. Schulte-Ostermann (Eds.), *Facets of Sociality* (pp. 153–168). Frankfurt: Ontos Verlag.

Gilbert, M. (2008). Two Approaches to Shared Intention: An Essay in the Philosophy of Social Phenomena. *Analyse & Kritik*, *30*(2), 483–514.

Gold, N. (2013). Team Reasoning, Framing, and Self-Control: An Aristotelian Account. In N. Levy (Ed.), *Addiction and Self-Control: Perspectives from Philosophy, Psychology, and Neuroscience* (pp. 48–66). Oxford, UK: Oxford University Press.

Gold, N. & Sugden, R. (2007). Collective Intentions and Team Agency. *Journal of Philosophy*, *104*(3), 109–37.

Hakli, R., Miller, K. & Tuomela, R. (2010). Two Kinds of We-Reasoning. *Economics and Philosophy, 26*(3), 291–320.

Hayek, F. A. (1960/2011). *The Constitution of Liberty: The Definitive Edition*. Chicago, IL: University of Chicago Press.

Hogg, M. A. & Abrams, D. (1998). *Social Identifications: A Social Psychology of Intergroup Relations and Group Processes*. London, UK: Routledge.

Hurley, S. (2005). Social Heuristics That Make Us Smarter. *Philosophical Psychology, 18*(5), 585–612.

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

Ludwig, K. (2016). *From Individual to Plural Agency. Collective Action: Volume 1*. Oxford, UK: Oxford University Press.

Mascaro, O. & Csibra, G. (2014). Infants Compute the Efficiency of Joint Actions. Poster presented at *The Fourth Annual Budapest CEU Conference on Cognitive Development*, Budapest, Hungary. Abstract retrieved on October 3, 2019 from <http://www.bcccd.org/down/bcccd_final_program_2014.pdf>.

Pacherie, E. (2011). Framing Joint Action. *Review of Philosophy and Psychology, 2*(2), 173–92.

Pacherie, E. (2013). Intentional Joint Agency: Shared Intention Lite. *Synthese*, *190*(10), 1817–39.

Petersson, B. (2007). Collectivity and Circularity. *Journal of Philosophy*, 104(3), 138–56.

Roth, A. S. (2014). Team Reasoning and Shared Intention. In A. Konzelmann Ziv & H. B. Schmid (Eds.), *Institutions, Emotions, and Group Agents: Contributions to Social Ontology* (pp. 279–295). Dordrecht: Springer Netherlands.

Searle, J. R. (1980). The Intentionality of Intention and Action. *Cognitive Science*, *4*(1), 47–70.

Sugden, R. (2000). Team Preferences. *Economics and Philosophy, 16*(2), 175–204.

Sugden, R. (2003). The Logic of Team Reasoning. *Philosophical Explorations, 6*(3), 165–81.

Török, G., Pomiechowska, B., Csibra, G., & Sebanz, N. (2019). Rationality in Joint Action: Maximizing Coefficiency in Coordination. *Psychological Science*, *30*(6), 930–941.

Tuomela, R. (2007). *The Philosophy of Sociality: The Shared Point of View*. New York, NY: Oxford University Press.