

Chapter 8

Modest Sociality, Minimal Cooperation and Natural Intersubjectivity



Michael Wilby

Abstract What is the relation between small-scale collaborative plans and the execution of those plans within interactive contexts? I argue here that joint attention has a key role in explaining how shared plans and shared intentions are executed in interactive contexts. Within singular action, attention plays the functional role of enabling intentional action to be guided by a prior intention. Within interactive joint action, it is joint attention, I argue, that plays a similar functional role of enabling the agents to act in a collaborative way such that their actions are rationally guided by a prior shared intention. This understanding of joint attention – as having a key functional role of enabling the rational guidance of joint intentional action by a prior shared intention – allows for an alternative understanding of the kind of minimal cooperation that infants can engage in. On this understanding, infants’ capacity to engage in joint actions is already an incipient capacity to engage in rational and intentional joint actions, albeit a capacity that is necessarily scaffolded by an adult rational co-partner.

Keywords Modest sociality · Minimal cooperation · Michael Bratman · Natural pedagogy · Shared intentions · Collective intentionality · Joint action · Joint attention · Natural intersubjectivity.

Joint action and shared collaborative activity can take place in a multitude of ways. There are long-term institutionalised collaborations – such as the construction of the International Space Station – occurring within a hierarchical structure, containing defined roles, interchangeable agents, and taking place over vast stretches of space and time. At a more modest level, there are non-institutionalised, not necessarily hierarchically-based, small-scale collaborative plans, such as going on a family trip, or hosting a surprise birthday party. At an even more fine-grained level, there are

M. Wilby (✉)

School of Humanities and Social Sciences, Anglia Ruskin University, Cambridge, UK
e-mail: michael.wilby@anglia.ac.uk

© Springer Nature Switzerland AG 2020

A. Fiebich (ed.), *Minimal Cooperation and Shared Agency*, Studies in the Philosophy of Sociality 11, https://doi.org/10.1007/978-3-030-29783-1_8

127

on-the-hoof, face-to-face interactions, such as taking a walk together, or moving an obstacle out of the way of the road.

In a way that runs partly parallel to this, the literature on Collective Intentionality and Shared Agency has tended to focus on three families of collaborative activities:

- (a) Massively Shared Agency
- (b) Modest Sociality
- (c) Minimal Cooperation

The first has cross-over with topics in *political philosophy*, the second with topics in *philosophy of mind*, and the third with topics in the *philosophy of cognitive science*.¹

It is not obvious, or perhaps even plausible, to suppose that there might be one overriding framework to capture all of these and related phenomena. Yet there remain questions of how they relate to each other. On the face of it, in at least some central cases, massively shared agency would not be possible without forms of modest sociality, and modest sociality would not be possible without forms of minimal cooperation. So the different areas must interface with each other in some places and at some points.

My concern here will be with the relation between (b) and (c). In particular, rather than discussing Minimal Cooperation as an alternative type of cooperation to Modest Sociality, I want to discuss the form that Minimal Cooperation might take when viewed as a way of executing shared (prior) intentions that arise as elements of plans within Modest Sociality (Bratman 1999, 2014). Approaching Minimal Cooperation in this manner, will, I think, allow for a different emphasis on what is required of the type of cooperation that humans might engage in at the face-to-face level. This kind of project differs in emphasis from the projects that have tended to motivate accounts of Minimal Cooperation.² The explicit motivation for such minimal accounts has tended to point to the discrepancy between childrens' capacity for shared agency, on the one hand, and the cognitively-heavy demands that accounts of Modest Sociality put on the protagonists, on the other. That is: infants can engage in collaborative activities, but they don't seem to have the sophisticated metarepresentational capacities that theories of Modest Sociality typically require agents to have for shared agency (Tollefsen 2005; Pacherie 2011; Butterfill 2012; Fantasia et al. 2014). The apparent discrepancy has led some to argue either that (a) Minimal Cooperation is to be applied to a distinct *explanandum* to that to which Modest Sociality applies – i.e., that collaborative infants are engaging in something other than genuine shared agency; or that (b) Minimal Cooperation is an alternative

¹ See, for instance, the collection of articles in Chant et al. (2014) which cover a range of ground. The phrase 'massively shared agency' comes from Shapiro (2014).

² Although (see Pacherie 2013 and Tollefsen 2014), for approaches similar to the one I shall be adopting. The question of how prior intentions can 'interface' with motor representations in the individual case is a topic that is beginning to receive quite a lot of attention, see in particular Butterfill and Sinigaglia (2014), Mylopoulos and Pacherie (2017), Brozzo (2017), Brownstein (2014), and Shepherd (2017).

explanans for the same phenomena that Modest Sociality is tackling – i.e., that infants do engage in genuine shared agency, but it should be modelled in an alternative way to the way in which it is modelled in the adult case (see Butterfil and Sebanz 2011 for a summary of these alternatives). There might well be room for both of these views – collaborative activities come in a range of flavours after all – and neither need be seen as competitors to accounts of Modest Sociality, which, on Bratman’s view at least, is multiply realisable.

My aim here is slightly different. I want to question what kind of account of Minimal Cooperation would be required for agents who *can* engage in the more sophisticated form of Modest Sociality, in a way that makes rational contact with the plans, conceptions and schemes that such agents are involved in. I shall argue that by taking such an approach we will find that we can understand forms of minimal cooperation that are at once suitable for those agents who see those minimal collaborations as elements within larger plans (i.e. adults), and for those agents who are not yet capable of seeing their collaborations in such terms (i.e. infants and young children). This is a useful endeavour because: (a) much – although not all – of the evidence that relates to infant social capacities is in their interactions with adults; and (b) it gives a glimpse into the question of how infants might be initiated into more sophisticated normative commitments and reason-giving that comes with engagement in larger plans and activities.

This chapter will be split into three parts: Sect. 8.1 shall outline an understanding of shared intentions within Modest Sociality, based on Michael Bratman’s work. Section 8.2 shall outline the notion of (what I shall call) *Natural Intersubjectivity*, which regards joint attention as having a pivotal role in coordinating and explaining small-scale interactions – of the kind that Minimal Cooperation is often used to explain – as manifestations of shared intentions. Section 8.3 shall argue that this notion of Natural Intersubjectivity is able to say something about the development of young children’s sociality, enculturation and mindreading skills.

8.1 Modest Sociality

The term *Modest Sociality* refers to shared agency – i.e. shared intentions and their attendant cooperative actions – within small adult groups, in the absence of asymmetric, hierarchical authority relations that typically constitute social relations within institutional structures. As Bratman puts it, the focus of Modest Sociality is:

primarily with duets and quartets rather than symphony orchestras with conductors, with small teams of builders rather than large and hierarchically structured companies, with small and informal neighbourhood groups rather than county governments, with small group discussion rather than deliberations in the US Senate, and with friendship and love rather than legally constituted marriage (Bratman 2014: 7).

Bratman distils his account of Modest Sociality³ into the following elements, which he states are “somewhat compressed sufficient conditions for our shared intention to J” (Bratman 2014: 103)⁴:

- (a) **Intention condition:** We each have intentions that we J; and we each intend that we J by way of each of our intentions that we J ... and by way of relevant mutual responsiveness in sub-plan and action, and so by way of sub-plans that mesh.
- (b) **Belief condition:** We each believe that if the intentions of each in favour of our J-ing persist, we will J by way of those intentions and relevant mutual responsiveness in sub-plan and action; and we each believe that there is interdependence in persistence of those intentions of each in favour of our J-ing.
- (c) **Interdependence condition:** There is interdependence in persistence of the intentions of each in favour of our J-ing.
- (d) **Common knowledge condition:** It is common knowledge that a–d.
- (e) **Mutual responsiveness condition:** Our shared intention to J leads to our J-ing by way of public mutual responsiveness in sub-intention and action that tracks the end intended by each of the joint activity by way of intentions of each in favour of that joint activity.

(a)–(d) relate to shared intentions, typically held prior to the joint action which they motivate, while (e) relates to the execution of those shared intentions within joint action. The core element to focus on here is the *Intention Condition*. The Belief Condition, Interdependence Condition and Common Knowledge Condition serve as ways of stabilising collaborations, to ensure smooth running, trust, mutual deliberation and bargaining. I shall turn to the *Mutual Responsiveness Condition* in Sect. 8.1.2, since this is where the account makes most obvious contact with Minimal Cooperation.

Consider an example. Suppose that we intend to dance the tango together. What does such an intention involve? The account parses this intention as follows: we each have an intention that we dance the tango, and we each intend that we dance the tango by way of each of those very intentions, as well as by way of relevant meshing sub-plans and actions. What makes the shared intention to dance the tango together different from us each having our own private intention to dance with the other is that we each intend that the dance come about by way of each of our respective intentions (and our respective and meshing sub-plans and actions by which we will realise this action). If all goes to plan, we will end with a joint action that comes about by way of each of our respective intentions – and meshing sub-plans – to perform that joint action.

Bratman’s approach is, as he calls it, *Constructivist*. Constructivism can be seen as a general methodological approach to understanding complex phenomena that constructs those complex phenomena out of simpler building blocks (Grice 1974). In the case of social agency, the building blocks are features of individual agency that combine to generate social phenomena. This is achieved, on Bratman’s account,

³There are a number of rival accounts of Modest Sociality (most notably Gilbert 2009; Searle 1990; Tuomela 2007), but I shall focus here on Bratman’s account at least partly because of its starting assumption of constructivism.

⁴The non-distilled accounts can be found in Bratman (1992), and then extended in *Shared Agency* (esp Chap. 4).

by putting the weight of the analysis on the *content* of the intentions. The intentions are standard, individual intentions. Within social contexts, they take on the same functional and normative load, as they would do within individual contexts. They are distinctively social only insofar as (a) two agents are involved; and (b) the content of the agents' intentions make reference to a joint activity and to the other agent's intentions with regards to that joint activity.

In order to do its Constructivist job there is already much that is contained within the *form* of an intention on Bratman's account. I shall discuss this now in a bit more detail, for, in doing so, we should be able to get a better grip on the challenges involved in working out how Bratmanian Modest Sociality can join-up with a more Minimal Coordination.

8.1.1 *The Planning Theory of Intention*

The central building block of Bratmanian Modest Sociality is the *Planning Theory of Intention* (Bratman 1987, 2007) that applies in the first instance to individuals. It resists the traditional idea that intentions can be understood reductively in terms of clusters of beliefs and desires. Rather, intentions are distinctive states of mind that are "typically elements within larger plans" (Bratman 1987: 28). Their functional role is to help coordinate our activities (both personal and social) in a way that compensates for our limited cognitive resources of memory, time and processing power. I shall outline some of the key feature of this theory.

Firstly, *plans are hierarchical and partial*. If one has an intention to give a talk in a far-away country, then that intention will be a core element of the plan that it generates. The plan will be *hierarchical*: the intention generates and embeds further sub-intentions (buying a ticket, booking accommodation, writing the talk), all of which can be decomposed into further embedded sub-sub-intentions. The plan is also *partial*: if I am to give a talk at the conference, then at some point I will need to find my way from my accommodation to the lecture hall – but that part of the plan can be left to later; I don't need to fill in all the elements in advance.

Secondly, *plans are stable in ends and more flexible in means*. By intending to give a talk in a far-away country, I thereby settle on a course of action and start thinking of a partial plan of how to fulfil it. The intention is stable in that it provides a settled objective, around which one draws up one's plans. Once it is in place, then there is a presumption of inertia that one will carry out the intention. This stability of ends allows for a flexibility of means – the lower down the hierarchy one goes the more flexibility one will allow in one's sub-intentions, so as to fulfil one's more robust core intention (e.g. one might intend to take the 10.00 am flight, but if that flight is not available, then, all else being equal, one will book a later flight, rather than just abandoning one's intention to give a talk).

Thirdly, *plans are subject to norms of consistency*. Plans, and the elements of which they are composed, must be coherent, both internally and epistemically. For instance, one shouldn't intend that one give a talk at the same time that one is

intending to visit a sick relative. Nor should one intend to give a public talk at a place where one believes no-one will be. When one buys a plane ticket one should ensure that it goes to the place where one intends to give the talk, and one should make sure it goes from an airport which is most appropriate for one to reach. All of this will require an internal means-end coherence between intentions and sub-intentions; consistency across intentions and across beliefs, as well as consistency in the agglomeration of plans, intentions and beliefs.

Fourthly, *intentions are motivational*. Intentions, in addition to the structural role mentioned above, provide agents with a conative, pro-attitude towards acting in a way that fulfils their content.

As can be seen from this, the emphasis on intentions as typically being elements of plans is an emphasis on *future-directed intentions*. This opens up the question of how intentions – as elements of partial plans – are connected to *intentional actions*. In order to approach this question, I want to make use of a three-way distinction, adapted from Elisabeth Pacherie (2006, 2011).

We can distinguish between: (i) distal intentions; (ii) proximal intentions; (iii) intentions in action.

The first are future-directed intentions, with a stability of content, subject to norms of consistency, and that generate, and are elements of, small-scale partial plans. We can note three aspects of these kinds of intentions: (a) They are states of intention, rather than intentional actions; (b) they tend to specify *types of outcomes* – e.g. that I present a talk in Geneva – rather than *token actions*; (c) because they specify general *outcomes* – states of affairs – rather than specific actions, then they can naturally be described with a ‘*that clause*’.⁵

The second are intentions that are directed towards particular token actions or token outcomes at a particular time – e.g. that I start speaking *now*. They will tend to be sub-intentions (or sub-sub-intentions, etc.) and so will tend to be more flexible and context-driven. We can note three aspects of these kinds of intentions: (a) They are states of intention, rather than intentional actions; (b) they tend to specify particular, token actions, rather than generalised outcomes; (c) when they specify particular actions, then they are most naturally described using the grammatical construction of ‘intention to’.

The third are the execution of an intention. We can note three aspects of these kinds of intentions: (a) they are intentional actions, rather than states of intention; (b) because they are actions, then they involve token actions rather than types of actions or generalised states of affairs; (c) because they are actions, then they are most naturally described using the grammatical adverb ‘intentionally’.

⁵The grammatical object of an ‘*intention that*’ takes a propositional form, rather than an agential or infinitival form (see Ferrero 2013). Consequently, an ‘intention that’ needn’t (but could) be self-reflexive: One could intend, for instance, that a parcel be delivered (if one were a manager of a postal company, for example), without there being any suggestion that it is oneself who delivers the parcel. By contrast, an *intention to* deliver a parcel – where the grammatical object of such an intention is an action, with a hidden but implied reflexive pronoun (PRO) – implies that the action is necessarily carried out by the subject of the sentence.

This is a very rough sketch of a set of important distinctions within action theory, and there is certainly room for both dispute and development with how I have set them out. Before turning to the ‘Interface Problem’ of how Future-Directed Intentions relate to Intentional Actions, I want to briefly say something about how the Planning Theory relates to shared intentions.

8.1.2 *Shared Intentions and the Planning Theory*

The Planning Theory tells us that there is more to having a shared intention than we J than merely two agents having volitional mental states with interlocking contents. It tells us, rather, that the agents have a *shared plan*.

In particular, it is worth making a distinction – which can be easily blurred – between a prior intention and a plan. Intentions are *elements* of plans. The plan itself is composed of a number of intentions which relate to each other in a hierarchical and coherent manner. If we think of a plan as something like a hierarchically structured ‘to-do list’, then there is a sense in which agents with shared intentions (which are in fact two intentions that interlock) genuinely *share* a plan (as a singular thing). Without a sense that there is one plan in place, then the constraints of coherence and consistency will get lost. The agents in question might possess different *copies* of this plan (different vehicles of the plan exist in the different agents heads), and the copies will likely contain different details which are relevant to playing ‘their part’, but, if there is to be the interconnectedness that is required – of coherence, consistency and agglomeration – then there is a sense in which there is a singular plan that the agents are both consulting. (It might be that the Common Knowledge Condition in Bratman’s condensed outline effectively plays the role of making this plan public, but I shan’t discuss that here).

8.1.3 *Shared Intentions and Mutual Responsiveness*

We have established the following set of distinctions between three faces of intentions (Table 8.1):

If the above is correct, then, in the execution of a prior intention (which might be dispersed and discontinuous over time), we move from the general to the specific, from intended outcomes to executed actions, and from an intention *that* to an

Table 8.1 Three faces of intention

Future-directed	Present-directed	Intentional action
Intention that p	Intention to φ	φ -ing intentionally
General	Specific	Specific
Directed at outcomes	Directed at actions	Actions

intention *to*. In Bratman's account, this is mentioned only insofar as this condition obtains:

- (e) **Mutual Responsiveness Condition:** Our shared intention to *J* leads to our *J*-ing by way of public mutual responsiveness in sub-intention and action that tracks the end intended by each of the joint activity by way of intentions of each in favour of that joint activity.

Although this doesn't tell us how the aspects are interfaced, it is useful to consider, since it contains within it all three faces of (shared) intention: the shared future-directed intention; the shared present-directed intention; and the shared intentional activity (Tollefsen 2014). That is to say, 'our shared intention to *J*' can be understood as a shared future-directed intention. Our 'sub-intentions' can be understood as being – depending on context – either future-directed sub-intentions, or present-directed intentions, and 'our *J*-ing' can be understood as a shared intentional activity.

So, we can parse the Mutual Responsiveness Condition as follows: our each intending that *p* (and our each intending that *p* by way of both those intentions and by way of meshing present-directed intentions and meshing actions, etc.), leads to our jointly intentionally *φ*-ing, by way of the 'public mutual responsiveness' of each of our respective future-directed sub-intentions that *q* and our present-directed intentions to *φ* that track the future-directed intentions that *p*.

For instance, suppose that we intend that we dance the tango for the next dance, and that this is common knowledge, in that it is something we have openly agreed to do. How is this shared intention executed? According to this account, we each intend the following: <that we dance the tango by way of each of our intentions to do so, as well as by our meshing future-directed sub-plans, our meshing present-directed sub-intentions, and our meshing actions>. This is the future-directed shared intention. The claim then is that this future-directed intention is executed because it 'leads to' our intentionally dancing *this very dance we are dancing together now*, and it led to this by way of a 'public mutual responsiveness' in both our future-directed sub-plans (that I hold out my hand, and that you stand up), and our present-directed intentions to dance *this very dance we are dancing* (the lead dancer intentionally moving as so, and the follower moving as thus), where these present-directed intentions (as well as the future-directed sub-intentions) track the initial future-directed intention (the initial agreement) that we dance the tango together.

The account claims that future-directed intentions 'lead to' intentional actions *by way of* present-directed intentions that *track* the future-directed intentions. So, the relation between a future-directed intention and an intentional action is not purely causal. One way to understand how this tracking exhibits itself in the linguistic expression of the intentional action is as an instance of what GEM Anscombe (1957) calls 'the intention with which a thing is done'. So: an intentional action to *φ* is done with the intention that *p*.⁶ That linguistically exhibits, but does not yet explain, the relation that holds between a future-directed intention and an intentional action.

⁶Where there is a future-directed sub-intention that *q* involved, one might just as truly describe it as an intention to *φ* with the intention that *q*. For instance, if we have a future-directed intention

To conclude this section: I have argued that a joint intentional action can, in some cases, be understood in terms of being an executed element of part of a shared prior plan. When it is understood in this way, we can describe the action as <jointly φ -ing with the intention that p >, where p is the content of the governing intention of the plan (and/or sub-plan). For instance, we might be dancing this very dance (a token instance of the tango) with the sub-intention that we dance the tango, and the fuller prior intention that our guests have an enjoyable evening. Here we have a prior shared intention (that our guests have an enjoyable evening) that leads to the joint intentional action, which is carried out by way of an intention to dance this very dance in a way that tracks both the sub-intention and the fuller prior intention. In a hierarchical way, both the sub-intention and, higher up the hierarchy, the prior intention, continue to shape and direct the execution of the action.⁷

8.2 Natural Intersubjectivity

How are prior shared intentions able to shape and direct the execution of a close-encounter joint action such that the action normatively tracks that prior shared intention? The claim that I shall pursue here is that it is via *joint attention* (or, at least, that joint attention has a pivotal role to play). Joint attention, I argue, functions to coordinate multiple agent's actions in a way that allows for the rational execution of a prior shared intention.⁸ This coordination allows for the agents' prior shared intention to guide the joint intentional action in a way that is under the agents' reliable conscious control. By jointly attending to the object, event or activity, the agents are able to monitor, control and respond to the ongoing activity, including the roles that each agent is playing as well as the salient features of the situation – such as the objects of their attention – in a way that normatively tracks, because rationally guided by, the shared prior intention.

This conception of a key functional role of attention as a matter of *conscious rational guidance of action* is related to various programmes within the Philosophy

that we hold a party, that might involve sub-plans that we decorate the room we hire in a colourful way. This in turn might lead to the jointly intentionally hanging *this* decoration *here*. This action is done intentionally. And we can describe it as being doing with the intention that we hold a party. And/or we can describe it as done with the intention that we decorate the room colourfully.

⁷Recent discussion of the relation between prior intentions and intentions-in-action has centred around a two-tier framework, in which a more primitive minimalist system runs partly independently of a slower, more purposeful, cognitive system that is conscious and propositional. This discussion has involved, inter alia, an account of how primitive agents (animals or infants) might be able to engage in joint actions via shared goals in the absence of the former (Butterfill 2012, 2013); and how the more complex system-2 might interface with the more primitive system-1 (Butterfill and Sinigaglia 2014; Sinigaglia and Butterfill 2015). See Christensen and Michael (2016) for a critical discussion of this proposal.

⁸Fiebich and Gallagher (2013) argue for the related idea that Joint Attention has the functional role of “[reducing] the risk of shared intention failing” (2013: 586).

of Attention (Campbell 2002; Wu 2011; Smithies 2011; Watzl 2017). I shall begin with a discussion of the relation between attention and action in the individual case, and then turn to a discussion of how this plays out in the joint case.

8.2.1 *Attention and Rational Guidance*

Attention is not sufficient for attentive action. Merely attending to an object or a scene is not going to necessitate acting on it (although it might raise urges or suggest specific actions). If attention is to have the coordinating and guiding role that I am claiming of it, then there needs to be something additional to perceptual attention involved; there must also be a conative aspect that combines with the cognitive aspect of attention. The additional conative aspect, I suggest, comes in through the distal, governing prior intention – when one is acting attentively, then one’s actions are under the normative *guidance* of one’s higher-order intentions/plans by means of one’s attention, where such intentions have a motivational force that carries through into the executed actions. By extension – in a manner that I shall develop in Sect. 8.2.3 – joint intentional action comes to be under the normative guidance of prior planned shared intentions by means of joint attention.

There are three aspects of the concept of guidance which are of relevance here.

Firstly, guidance is something achieved by a person rather than by a sub-process, or “independent causal mechanism”, such as a motor representation (Frankfurt 1978). On the assumption that higher-order plans have “agential authority” that amounts to a form of ‘self-governance’ (Bratman 2007), then it is these higher-order plans that are doing the guiding.

Secondly, guidance is not the same as merely *causally effecting an outcome*. Guidance involves a personal level awareness of what one is doing and needs to do; how one is to do it; and of the circumstances that one is in, that would allow one to intervene, correct and control one’s actions as and when necessary in order to fulfil one’s plans and intentions. Guidance will involve exercising one’s conscious control at some moments, submitting to the unreflective, skilled ‘flow’ (Csíkszentmihályi 1975; 1990; Schear 2013) of action at others, and consciously intervening in or correcting those actions at still others. When one is unreflectively engaged in the flow of a skilled activity – as a well-practised pianist might be – then the activity remains under the guidance of one’s plans and intentions, such as the intention to play a particular piece at a certain tempo (Velleman 2008). The skilled pianist will be ready to intervene and correct slips, and be ready to take more active control at certain points of certain sequences of fine-grained motor control that might at other times be performed automatically (see Brozzo 2017).

Attention is a capacity by which one can make rationally informed decisions about how or when to control, intervene, correct, or ‘go with the flow’ in pursuit of a goal or plan: it is by means of attention that one is able to determine what skills to put into action and when, and to which objects or locations one should direct one’s

skills towards. If one was not attending to the scene that one was in – the activity itself as well as the surroundings – then one would have immense difficulty exercising one’s capacities successfully. One’s attention can be directed in a ‘top-down’ manner, as when one is guiding one’s actions in a controlled and deliberative way; or it can be grabbed in a ‘bottom-up’ manner by unexpected or unanticipated events in the environment, allowing one to intervene or correct when one is acting ‘automatically’.⁹ Such unexpected events might require one to change or abort one’s more flexible sub-intentions in pursuit of a more robust distal intention, or even to unexpectedly achieve one’s intentions and plans by capitalising on a lucky accident. The notion of ‘guidance’ allows for unanticipated interruptions – events that grab or disrupt one’s attention – to nevertheless have a place within a rational and planned intentional schema.

Thirdly, guidance sees an intentional action as an actualisation of a prior intention. When one’s intentional actions are carried out under the rational guidance of one’s prior intention, then we can say that the action is carried out with the intention that p , where p is the content of the prior intention.¹⁰ When one is acting intentionally with the intention that p , then there is a sense in which the planned intention is operative *in* the action, rather than being a mere causal precursor to the action. The planned intention, in having ‘agential ‘self-governing’ authority’, has guidance control of a range of one’s sub-systems and bodily movements. Intentional actions are thereby “exercises of conceptual capacities”, to use McDowell’s (2009) phrase.¹¹

8.2.2 Attention and Demonstrative Reference

I have suggested that attention plays a role in exercising the rational guidance of intentional action by a prior intention. As we shall discuss in the next section, in the case of joint attention, we can say that joint attention takes the role of *rationally coordinating* joint intentional actions such that they are exercises of, because rationally guided by, a shared prior intention. The question I want to discuss first, in this section, is how attention *simpliciter* realises the broad functional role I have sketched.

A key functional role of attention, then, is to enable intentional actions such that they are rationally guided by a higher-order intention and plan. How does it do this? How is this role *realised*? Attention can be seen to be a form of demonstrative

⁹Or one can make use of a combination of these and guide one’s attention in a ‘top-down’ manner in order to find something that will grab one’s attention and ‘pop-out’ from the scene.

¹⁰And the action can be intentional under more than one description (Davidson 1963), depending on the plan/sub-plan that we choose to highlight.

¹¹Luthra (2016) calls this kind of view *Rationalism about agential control*: whereby “our ability to determine, as agents, how we act consists solely in guidance of our actions through the exercise of those sorts of rational capacities—capacities for intention, practical judgment, practical reasoning, and the like” (Luthra 2016: 2272)

reference (Campbell 2002; Dickie 2015) – a way of specifying particular objects, processes, features, locations or events within one’s perceptual field, that makes them accessible to thought; ‘this thing over here’ and ‘that thing over there’. In doing this, attention can act as a kind of interface between world, thought and action (Campbell 2002, Ch 7). As mentioned earlier, prior intentions will, by their nature, be more general than the specific actions that actualise those intentions. Attention allows the general to be linked to the particular by identifying particular objects, processes, features, locations or events that can act as determinants of the general concepts that feature within elements of plans and sub-plans (see Evans 1982, and the notion of a fundamental ground of thought, for an outline of how this connection might be made).

Suppose, for instance, that one intends that one have baked beans for dinner tonight. One needs, in that case, to be able to identify a can of baked beans when one sees one. Further, the sub-plans, as and when they come about and get filled in, will involve further general concepts that need to be identified and realised by particulars in one’s environment (pans, spoons, plates, etc.), as well as particular actions (reaching, grabbing, etc.) and spatial relations (to the right, up the top, to the left, etc.). This requires the agent to be able to have demonstrative thoughts about these objects, relations and actions, in a way that will link up with the plans. Those demonstrative thoughts are grounded in attention – it is by means of attention that one can identify *that object* as a can of baked beans, and *that object* as a pan, and then, when it comes to action, calculate the distance that needs to be traversed between them (*that distance*), and then perhaps, by means of indexical predicates, identify the needed movement and actions: I should do *that action* (Heal 1997; Butterfill and Sinigaglia 2014).¹²

An intention that specifies, for instance, <that I cook baked beans tonight>, gets actualised partly by means of a demonstrative concept grounded in attention that specifies <that *those* are baked beans>. One can then insert that specification into the generalised plan, so as to intend the more specific plan <that I cook *those* beans>, eventuating, via numerous further, finer-grained sub-plans, involving pots and pans, in an intention to cook the beans, and the very act of intentionally cooking the beans. On a hierarchical planning theory, such intentions will be filled out with particulars – demonstratives, indexical predicates, etc. – all the way down, as sub-plans are filled in. At each stage the plans and sub-plans are guiding the actions by means of the attentive reference to places, locations, features, etc., that fulfil the plans. The demonstratives have the role of allowing the agent to insert objects, actions, properties and locations into a propositional level, planning schema.

To summarise this and the preceding section: I have argued that attention plays the role of enabling the rational, guided execution of intentional actions by prior intentions and prior plans. It plays this role partly by means of demonstratively identifying the specific features of the environment – objects, locations, processes,

¹²Although much of the philosophical and psychological literature on attention has been concerned with object attention, I am assuming here that one can likewise attend to processes and actions and events.

actions – that can act as particular instantiations of the elements that figure in prior plans in only a general way. One then acts on those objects, locations, processes and actions in accordance with what would be required for the fulfilment of the plans, in a way that is guided by those plans, and not merely caused by them. On this picture, motor representations play a role as the sub-personal mechanisms by which this guidance is carried out, but not as intentions or even intention substitutes (cf, Butterfill and Sinigaglia 2014).

8.2.3 *Joint Attention and Coordination*

The above sketches a role for attention within action, whereby attention is the means by which plans and sub-plans can guide intentional actions. I now want to argue that *joint attention* plays a corresponding role in interactive cases of joint agency, where part of that corresponding role is to act as a means of *rationaly coordinating* joint action in accordance with, because guided by, a prior shared intention.

It is important here to fix the kind of situations I am talking about. There are of course ways of executing shared intentions which don't require interaction. If we are searching for a lost dog, for instance, and you go north and I go south, then there is no interaction, even though there is a shared intention that the dog be found by one of us, and various sub-intentions to be executed in accordance with that plan. My concern, rather, is with interactional contexts with coordinated actions. These will be *cooperatively loaded* actions – actions where coordination (and not merely 'meshing') is essential to their fulfilment.

'Joint attention', as I shall be using the term, refers to a situation where two or more agents are both attending to an event, object, process, feature or location, and *they are aware of this whole awareness*,¹³ which is to say that their attention, and its common content, is 'out in the open' and 'mutually manifest' to each.¹⁴ Attention (and by extension joint attention) here is to be understood not as a full-blown propositional attitude, with a 'that'-clause and an independently assessable truth-value, but rather as an unsaturated demonstrative element that can figure in a propositional attitude, such as a demonstrative singular thought, that can, in turn, act as the realisation of an element of a plan. For instance, attention to an object – a can of beans, say – can give specific content to a general intention <that one cook beans for dinner tonight>, by specifying the particular can of beans that can fulfil that intention.

Now, the claim at hand is that joint attention has the role of coordinating an interactive joint action in accordance with, because rational guided by, a shared prior

¹³This way of phrasing it comes from Peacocke (2005).

¹⁴This understanding of joint attention is a short-hand for a number of alternative ways of filling out the idea that might take a truncated, iterative, relational or fixed-point form (see Barwise 1988). I shall leave it unstated what form it should take, since any attempt to do so would require a full-length paper of its own, although my sympathies are with the relational approach (Campbell 2005; Eilan 2007; Seemann 2012; Wilby 2010).

intention. Suppose, to take John Searle's (1990) well known example, we intend that we make a Hollandaise Sauce together. We have a plan. We know how to carry it out. We know the ingredients needed and the recipe to follow. And we have an open agreement that we will make it this afternoon. When the time comes we will need to identify the objects, processes, angles, locations, etc., that will be necessary to carrying out this prior shared intention. We will need to coordinate on the details – we will need to work out who is pouring, who is stirring, where the bowl is, the timing, etc. What we attend to will help determine what we use and how we use it in a way that tracks the shared prior intention. That we are both attentive to the elements that comprise the situation we are in is clearly central to linking our activities and our plans. It is because we are both attending to the unfolding activities – by means of attending to the objects, locations, relations, processes and features of those activities – that those activities can be understood as *realisations* of our shared planning. And it is by understanding those activities as realisations of our shared planning that allows for the shared planning to rationally guide the ongoing activities. By moving the pan just so, and cracking the egg just so, we can, providing we have a hierarchical plan in place with sub-plans and sub-sub-plans and sub-sub-sub-plans (etc.), place those very activities into that hierarchical structure. Our respective attention to the ongoing situation and environment allows us to coordinate our actions so that those actions – both before and after execution – can be recognised as elements of a plan that we are executing together, and that we must track and update together.

One might object at this point that *joint* attention – in the full-blown ‘mutually manifest’ sense – is not needed. Each agent needs to fulfil their own part of the joint action, and so needs to attentively monitor what is necessary for playing that part. Perhaps that might mean monitoring the other's *actions* at times, and the results of those actions, but why need there be anything more substantial than that? There is no need here, it might be thought, for our respective perceptual attention to be *out in the open* in any robust sense. Perhaps all that is needed is parallel attention to the ongoing action.

In response to this objection, it is worth considering what I shall call the *guided rationality of joint agency*. Suppose we are engaging in a joint intentional action *with the shared intention that we make a Hollandaise Sauce*. For this joint action to be intentional under this description it must be being guided by that shared intention; or, at least, that has been the thrust of the argument thus far. This is to say – once decomposed into its constructivist Bratmanian elements – I must be being guided by *your* intention that we make Hollandaise Sause, just as much as I'm guided by mine. What would it mean for me to be guided by your intention? How can your intention get a grip on me such that it *guides* my actions? If we were to think of your intention guiding my actions in precisely the way that my intention sometimes *causally controls* mine, then we truly would be violating an ‘own action’ condition on intention in an implausible way. For it is not as if I can get inside your skin, so to speak, and make your limbs move.

One way your intention might be thought to rationally guide my actions would be for my actions to be visually guided by your observable limb movements and

their effects. I observe you shifting the pan this way, causing me to pour in that way. Given that your actions and movements are themselves guided by your intention that we make the Hollandaise Sauce, then, it might be argued, if we suppose that guidance is transitive, I will thereby be guided by your intention. But this isn't rational guidance. It is, rather, a matter of reactively responding to your guided intentions. Rational guidance, as I shall explain shortly, is something more than this.

Another way in which your intention might be thought to guide my actions would be if I were to act in a way that I thought was responsive to your wishes and intentions, such that I adopted your prior intentions as if they were my own. You shift the pan this way, causing me to pour in that way because I am concerned to fulfil what you had intended. But this also isn't rational guidance – your intentions are here normatively *influencing* my intentional actions (they provide me with a reason to act in a certain way perhaps), but they are not rationally guiding my actions. They are not rationally guiding my actions because your intention, in this case, is not poised to control, intervene, correct or respond to my actions in pursuit of executing the prior intention.

Guidance, then, is not merely having one's actions being causally affected by a prior event such as an intention (Frankfurt 1978), nor is it a matter of merely being rationally responsive to the intentions of another, as if satisfying another's intentions were a reason for my action. Rather, as already noted, rational guidance is a matter of having a multi-faceted capacity and skill to control, intervene in, respond to, or even 'go with the flow' with regards to, one's actions in pursuit of the end specified in the content of the prior intention, such that one is able to anticipate effects, intervene to prevent mishaps, or respond to correct errors. If I am to be partially *rationally guided* by your intentions (as aspects of a shared intention) then your intentions must be poised to partially control, intervene in, correct, step back from, or be responsive to, my actions in pursuit of our shared intention. That is to say: your intention <that we make a Hollandaise Sauce together by means of each of our intentions and meshing sub-plans> must be poised to have that kind of guiding influence over my actions.

The present suggestion is that joint attention – in its full-blown 'mutually manifest' guise – plays the functional role of enabling your intention, in combination with my intention, to have that guiding role over both your actions, and my actions, (which together constitute the joint action) by making mutually manifest the elements of the environment that contribute to fulfilling the corresponding elements of the prior plan of which our shared distal prior intention was the governing element. By making features of the environment mutually manifest, our actions will be directed, in meshing, complimentary ways, towards fulfilling the shared intentions and the shared plan. Without us being mutually aware of the relevant features of the environment in which we are operating (the objects, the processes, the actions and the 'shared', not purely ego-centric space in which this occurs), then we would be unable to update, adapt and consult the course of our shared plan – what needs to be done and what has been done – in the necessary way.

Part of what makes joint attention important here is that plans are partial and are flexible in means. As a consequence, the execution of some sub-intentions will be

on the fly. Suppose, for instance, that we have a partial plan involving numerous fairly fine-grained tasks that could be performed by either one of us when the time comes. If we are to operate broadly in synch with each other then we need to be aware of who has done what, who is doing what, and who should be doing what, when and how (given the present lay-out of the environment and present locations of the participants). The features of the environment dictate our actions insofar as they are features that will fulfil our plans and are located within our field of attention. Mere parallel attention will not allow us do this because, without knowing what the other agent is aware of, we will not be able to anticipate their actions; we will be unable to determine the layout of the environment that is available to them and at what stage of the shared plan they think we are at.

It might be thought that the agents need only have awareness of each other's attention to the environment, but not that it be 'mutually manifest'.¹⁵ But mutual manifestation has a role to play here which mere awareness of each other's attention would not fulfil. If we are *just* aware of each other's attention then, ex hypothesis, we are not aware of the other's awareness of our attention. That is to say, I might just think that you think that I am an inattentive fool. But this would play havoc with our activity. If a functional role of attention is to implement a multi-faceted skill of guidance, then presumed inattentiveness on either of our parts should lead to the other constantly intervening and correcting for our predicted bumbling, inattentive actions. Consequently, if I suspect that you think that I am an inattentive fool, then I should expect you to constantly interfere in my actions, and I will be compensating for that.

Joint attention – in its full-blown mutually manifest sense – allows us to be guided by a prior shared intention and plan. It is because we are *jointly* attending to the environment and the features in it, that your intention and my intention (which comprise in the appropriate Bratmanian sense our shared intention) together rationally guide the joint action. We are together attentively monitoring the on-going activity – what needs to be done, what has been done, what needs correcting, what needs intervention, what needs skilled, automatic activity – in a way that is open. This openness is crucial. It enables both of us to anticipate, intervene, control and respond to the ongoing activity *as a manifestation of a shared intention*, rather than merely reactively responding to the other. These are manifestations of a shared intention because they are actions that are fulfilling and forwarding what is required in order to fulfil that intention as an element of a hierarchically ordered, structured and coordinated shared plan that, in its details, will be flexible and constantly updated. It is not merely the ongoing action, or the relevant elements of the environment, that are perceptually open to us as a manifestation of a shared intention, but also, when they occur, our respective tendencies to correct and intervene (e.g. I tell you to bring the pan closer, or you tell me to 'watch out' as I am about to drop the sauce).

¹⁵Blomberg (2016) argues for a point related to this with regards to Common Knowledge and Joint Action

If we are both attentive to the objects, locations, actions and events in front of us – and know that we are both so attentive – then we can immediately conceptualise and frame the actions that are performed (be they controlled, intervening or correcting) as actions that are directed towards, because guided by, a prior shared intention. Joint attention, because it interfaces with prior intentions, allows us to *mutually recognise* both the actions which are being intervened on, as well as the intervening actions, *as manifestations of a prior shared intention*.

8.3 Natural Intersubjectivity and Development

I have argued that minimal cooperation can interface with shared intentions by way of joint attention. Joint attention allows for an ongoing joint activity to be rationally guided and coordinated by a prior shared intention. The claim is that in such cases, there is no need to appeal to a more minimal concept of a ‘shared goal’.

A question which opens up about this, however, is what to say about three kinds of cases where there would appear to be joint activity, but where shared intentions are not operative (Tollefsen 2014). These are: (a) spontaneous, ‘flash mob’ joint actions; (b) interactions involving young children and infants; (c) interactions between animals. I shall focus on infant interaction, although more would need to be said about the others.¹⁶

8.3.1 *The Development of Shared Intentions*

I do not have space to fully develop the claim but I shall argue that joint attention and joint action play a developmental role in allowing children to engage in structured activities that, although not actually governed by shared intentions, are nevertheless enacted in a way that partially assumes the existence of shared intentions and shared plans. The capacity for joint attention allows infants to conceptualise the joint activities that they are engaged in as having a place within a wider normative structure, even when they are not yet capable of fully grasping that normative structure or the content of some of the elements of that structure.

¹⁶Very briefly – I would argue that Flash Mob actions are defined in terms of their being self-consciously *unplanned*. To the extent that this is part of their motivation and definition, then they are defined and motivated (negatively) in terms of planning. With regards to animal interaction, I would follow Tomasello et al. (2005) – or perhaps more specifically, Henrike Moll’s *cultural transformation* version of that hypothesis (Moll 2016, 2018; Kern and Moll 2017) – that shared intentionality is a defining feature of human uniqueness, and so animal interaction, unlike infant interactive cognition, is to be understood in a minimal way, along the lines suggested by Apperly and Butterfill (2009).

It would be useful, I think, to begin with the metaphorical picture I sketched in Sect. 8.1 and that was operative in the background of the discussion in the previous section. We can think of plans as a kind of hierarchical diagram, with lines drawn from intentions to sub-intentions and from sub-intentions to sub-sub-intentions, and so on. As Bratman stresses, some sub-intentions will be bargained over (who does what when and how), so we can think of each of these boxes as being filled in with further content as and when the time comes, and boxes being added and taken away. Further, we can think of joint attention as allowing agents to (a) identify how to execute the sub-intention which appears in a sub-intention box, and then, (b) when the sub-intention has been executed, mentally ‘ticking’ the box, so as to mentally record its successful execution. Now, if, as I suggested before, shared plans are shared in a fairly robust way – they operate, functionally at least, like a wall-chart or online schedule which we can both edit – then we can see, as argued, in the previous section, why Joint Attention must be something more than parallel; it must be playing a public role to ensure that we are updating our copies of the plan together.

With this metaphor in hand, I want to suggest that infants and children often engage in activities with a sense that they are doing something as part of a larger, more structured prior shared plan, albeit a shared plan that they don’t fully understand or have access to. They might engage in games with rules for instance, yet will not know the rules beforehand, and learn them through engaging in the activities themselves (Wittgenstein 1956; Bruner 1993). Similarly, children might engage in more informal plans, but not have a full idea of what the plan is. They understand that what they are doing has a role, and they can begin to piece it together as they proceed through the activity. So, it is a kind of reverse picture of the planning theory: they might have a sense that there are ‘boxes’ to be ticked off and filled in, but not have an idea of how these boxes are structured, how they relate, or what their content is. Children will work from the activities to an understanding of the sub-plans to an understanding of the whole plan. In many cases, it might well be that the infants and children have an initial idea of a shared prior intention (‘that we make a cake together’), without yet having an understanding of either what it means for it to be genuinely shared, or what is required to fill it out. Their understanding of what is required for a plan to be shared comes about, on this picture, by means of engaging and ‘pitching in’ (Paradise and Rogoff 2009) in collaborative activities. Only when they have mastered these activities, and the way that actions join with sub-intentions and sub-intentions with intentions, will they begin to see what a robust shared intention is.

By learning smaller activities involved in baking cakes, for instance, children will, as they go along, be learning how to fit those activities into a wider plan, and will come to see the activity as an intentional activity with the intention that *p*. In coming to engage in the activities, they might also eventually come to have an understanding of the complex structure of a shared intention; that is, of what it is to share an intention such that the content of one’s own plan is intertwined with the content of the other’s plan, and where one is responsive to the wishes and plans of the other and how they might mesh with one’s own.

In some sense, this is illusory. There is no genuine *shared* intention between the caregiver and the infant. But the claim here is that the pedagogical attitude of a caregiver is (sometimes, in structured activities) one of a caregiver making a shared plan that is constructed *as if* it were equally the intention of the child, even when the child is not yet able to conceptualise the complex intentions that go into a shared intention, or the complex hierarchy by which it takes shape.

As with the adult case, joint attention plays a key role. By jointly attending, the participants are able to monitor, intervene, control and respond to the ongoing activity. At a very early stage, infants will almost always be the recipients of this guidance. The guidance will come almost fully from the caregiver's ersatz shared intention.

This view is related to the idea of *Natural Pedagogy* as developed by Gegerly and Csibra (2006), Csibra and Gegerly (2009, 2011), which postulates a “well-organised package of biases, tendencies and skills” (Gergely and Csibra 2006: 8) for learning on the side of children and infants, and a natural tendency towards structured teaching on the side of caregiver adults (Csibra and Gegerly 2011). Csibra and Gegerly's focus is on the often cognitively opaque *cultural knowledge* that natural pedagogy helps develop; knowledge which children often take to be generic, holding across contexts and people (Egyed et al. 2013; Topál et al. 2008), and which is often imparted within a formal context. This suggests a relation between Minimal Cooperation and Massively Shared Agency. The focus here, however, has been on the development of understanding others' plans and how those plans interact with one's own actions and how this might develop into a more sophisticated capacity for shared intentions. This is a relation between Minimal Cooperation and the contingent and local plans that arise in Modest Sociality, and is perhaps more closely related to what Barbara Rogoff calls ‘informal learning’ (Rogoff et al. 2016).

8.4 Concluding Remarks

To summarise, I have argued that we need an understanding of how Bratmanian shared intentions might interface with joint intentional actions, such that joint intentional actions can be seen as executions of shared intentions, which track those prior shared intentions in a way that would allow us to say that the agents are jointly φ -ing with the intention that p . That this is something of a lacuna in traditional accounts of collective intentionality, is not a new complaint (see, e.g. Tollefsen and Dale 2012, who call this the ‘execution problem’). But I have attempted to answer this question by arguing that joint attention plays a pivotal functional role in this by enabling the agents' actions to be coordinated in a way that is rationally guided by and responsive to their prior shared intentions. Finally, I argued that this might help provide a framework for explaining the development of interactive skills within an informal learning environment.

References

- Apperly, I., & Butterfill, S. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953–970.
- Anscombe, G. E. M. (1957). *Intention*. Cambridge, MA: Harvard University Press.
- Barwise, J. (1988). Three views of common knowledge. *Proceedings of the 2nd conference on theoretical aspects of reasoning about knowledge* (pp. 365–379).
- Blomberg, O. (2016). Common knowledge and reductionism about shared agency. *Australasian Journal of Philosophy*, *94*(2), 315–326.
- Bratman, M. (1987). *Intention, plans, and practical reason*. Cambridge, MA: Harvard University Press.
- Bratman, M. (1999). *Faces of intention: Selected essays on intention and agency*. Cambridge: Cambridge University Press.
- Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review*, *101*(2), 327–341.
- Bratman, M. E. (2007). *Structures of agency: Essays*. New York: Oxford University Press.
- Bratman, M. E. (2014). *Shared agency: A planning theory of acting together*. Oxford: Oxford University Press.
- Brownstein, M. (2014). Rationalizing flow: Agency in skilled unreflective action. *Philosophical Studies*, *168*(2), 545–568.
- Brozzo, C. (2017). Motor intentions: How intentions and motor representations come together. *Mind & Language*, *32*(2), 231–256.
- Bruner, J. (1993). *Act of meaning*. Cambridge, MA: Harvard University Press.
- Butterfill, S. A. (2012). Joint action and development. *The Philosophical Quarterly*, *62*(246), 23–47.
- Butterfill, S. A. (2013). Interacting Mindreaders. *Philosophical Studies*, *165*(3), 841–863.
- Butterfill, S. A., & Sebanz, N. (2011). Joint action: What is shared? *Review of Philosophy and Psychology*, *2*(2), 137–146.
- Butterfill, S. A., & Sinigaglia, C. (2014). Intention and motor representation in purposive action. *Philosophy and Phenomenological Research*, *88*(1), 119–145.
- Campbell, J. (2002). *Reference and consciousness*. New York: Oxford University Press.
- Campbell, J. (2005). Joint attention and common knowledge. In N. Eilan et al. (Eds.), *Joint attention: Communication and other minds* (pp. 287–297). Oxford: Oxford University Press.
- Chant, S. R., Hindriks, F., & Preyer, G. (Eds.). (2014). *From individual to collective intentionality*. Oxford: Oxford University Press.
- Christensen, W., & Michael, J. (2016). From two systems to a multi-systems architecture for mindreading. *New Ideas in Psychology*, *40*, 48–64.
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153.
- Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B*, *360*, 1149–1157.
- Csikszentmihályi, M. (1975). *Beyond boredom and anxiety*. San Francisco: Jossey-Bass Publishers.
- Csikszentmihályi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.
- Davidson, D. (1963). Actions, reasons, and causes. *The Journal of Philosophy*, *60*(23), 685–700.
- Dickie, I. (2015). *Fixing reference*. New York: Oxford University Press.
- Egyed, K., Kirly, I., & Gergely, G. (2013). Communicating shared knowledge in infancy. *Psychological Science*, *24*(7), 1348–1353.
- Eilan, N. (2007). Consciousness, self-consciousness and communication. In T. Baldwin (Ed.), *Reading Merleau-Ponty*. London: Routledge.
- Evans, G. (1982). In J. McDowell (Ed.), *The varieties of reference*. New York: Clarendon Press Oxford.
- Fantasia, V., De Jaeger, H., & Fasulo, A. (2014). We can work it out: An enactive look at cooperation. *Frontiers in Psychology*, *5*, 874.

- Ferrero, L. (2013). Can I only intend my own actions? *Oxford Studies in Agency and Responsibility*, 1, 70–94.
- Fiebich, A., & Gallagher, S. (2013). Joint attention in joint action. *Philosophical Psychology*, 26(4), 571–587.
- Frankfurt, H. G. (1978). The problem of action. *American Philosophical Quarterly*, 15(2), 157–162.
- Gergely, G., & Csibra, G. (2006). Sylvia's recipe: The role of imitation and pedagogy in the transmission of cultural knowledge. In: *Roots of human sociality: Culture, cognition, and human interactions* (pp. 229–255).
- Gilbert, M. (2009). Shared intention and personal intentions. *Philosophical Studies*, 144(1), 167–187.
- Grice, H. P. (1974). Method in philosophical psychology (from the Banal to the Bizarre). *Proceedings and Addresses of the American Philosophical Association*, 48, 25–53.
- Heal, J. (1997). Indexical predicates and their uses. *Mind*, 106(424), 619–640. Reprinted in J. Heal (2003) *Mind, Reason and Imagination*. Cambridge University Press.
- Kern, A., & Moll, H. (2017). On the transformative character of collective intentionality and the uniqueness of the human. *Philosophical Psychology*, 30(3), 315–333.
- Luthra, Y. (2016). Non-rational aspects of skilled agency. *Philosophical Studies*, 173(8), 2267–2289.
- McDowell, J. (2009). Acting as one intends. In J. Dancy & C. Sands (Eds.), *Philosophy of action: An anthology*. Oxford: Wiley-Blackwell.
- Moll, H. (2016). Tension in the natural history of human thinking. *Journal of Social Ontology*, 2(1), 65–73.
- Moll, H. (2018). The transformative cultural intelligence hypothesis: Evidence from young children's problem solving. *Review of Philosophy and Psychology*, 9(1), 161.
- Mylopoulos, M., & Pacherie, E. (2017). Intentions and motor representations: The interface challenge. *Review of Philosophy and Psychology*, 8(2), 317–336.
- Pacherie, E. (2006). Towards a dynamic theory of intentions. In: *Does consciousness cause behavior* (pp. 145–167).
- Pacherie, E. (2011). The phenomenology of joint action: Self-agency versus joint agency. In A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience*. Cambridge, MA: MIT Press.
- Pacherie, E. (2013). Intentional joint agency: Shared intention lite. *Synthese*, 190(10), 1817–1839.
- Paradise, R., & Rogoff, B. (2009). Side by side: Learning by observing and pitching in. *Ethos*, 37(1), 102–138.
- Peacocke, C. (2005). Joint attention: Its nature, reflexivity, and relation to common knowledge. In N. Eilan, C. Hoerl, T. McCormack, & J. Poessler (Eds.), *Joint attention: Communication and other minds: Issues in philosophy and psychology* (pp. 298–324). Oxford: Clarendon Press.
- Rogoff, B., Callanan, M., Gutierrez, K. D., & Erickson, F. (2016). The organization of informal learning. *Review of Research in Education*, 40(1), 356–401.
- Scheur, J. K. (Ed.). (2013). *Mind, reason, and being-in-the-world: The McDowell-Dreyfus debate*. New York: Routledge.
- Searle, J. (1990). Collective intentions and actions. In P. R. Cohen, J. Morgan, & M. Pollack (Eds.), *Intentions in communication* (pp. 401–415). Cambridge, MA: MIT Press.
- Seemann, A. (2012). Joint attention: Towards a relational account. In A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 183–202). Cambridge, MA/London: MIT Press.
- Shapiro, S. (2014). Massively shared agency. In M. Vargas & G. Yaffe (Eds.), *Rational and social agency: The philosophy of Michael Bratman* (pp. 257–293). Oxford: Oxford University Press.
- Shepherd, J. (2017). Skilled action and the double life of intention. *Philosophy and Phenomenological Research*, Online First.
- Sinigaglia, C., & Butterfill, S. A. (2015). On a puzzle about relations between thought, experience and the motoric. *Synthese*, 192(6), 1923–1936.

- Smithies, D. (2011). Attention is rational-access consciousness. In C. Mole, D. Smithies, & W. Wu (Eds.), *Attention: Philosophical and psychological essays* (pp. 247–273). New York: Oxford University Press.
- Tollefsen, D. (2014). A dynamic theory of shared intention and the phenomenology of joint action. In Chant et al. (Eds.), *From individual to collective intentionality: New essays* (pp. 13–33). New York: Oxford University Press.
- Tollefsen, D. (2005). Lets pretend! Joint action and young children. *Philosophy of the Social Sciences*, 35(1), 75–97.
- Tollefsen, D., & Dale, R. (2012). Naturalizing joint action: A process-based approach. *Philosophical Psychology*, 25(3), 385–407.
- Topál, J., Gergely, G., Miklósi, Á., Erdőhegyi, Á., & Csibra, G. (2008). Infants' perseverative search errors are induced by pragmatic misinterpretation. *Science*, 321(5897), 1831–1834.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioural and Brain Sciences*, 28, 675–735.
- Tuomela, R. (2007). *The philosophy of sociality: The shared point of view*. Oxford: Oxford University Press.
- Velleman, J. D. (2008). The way of the wanton. In C. Mackenzie & K. Atkins (Eds.), *Practical identity and narrative agency*. New York: Routledge.
- Watzl, S. (2017). *Structuring the mind: The nature of attention and how it shapes consciousness*. Oxford: Oxford University Press.
- Wilby, M. (2010). The simplicity of mutual knowledge. *Philosophical Explorations*, 13(2), 83–100.
- Wittgenstein, L. (2001) [1956]. *Philosophical investigations*. Oxford: Blackwell Publishing.
- Wu, W. (2011). Confronting many-many problems: Attention and agentive control. *Noûs*, 45(1), 50–76.