

Joint Attention and Communication*

Rory Harder

January 12, 2022

Abstract

Joint attention occurs when two (or more) individuals attend together to some object. It has been identified by psychologists as an early form of our joint engagement, and is thought to provide us with an understanding of other minds that is basic in that sophisticated conceptual resources are not involved. Accordingly, it has also attracted the interest of philosophers. Moreover, a very recent trend in the psychological and philosophical literature on joint attention consists of developing the suggestion that it holds partially in virtue of communication: it is because we share our thoughts or feelings about an object that our individual attention becomes joint. This paper unpacks the communicative suggestion in a way that accounts for joint attention's basicness.

Keywords: Joint Attention, Joint Action, Griceanism, Common Knowledge, Theory of Mind

Contents

o	Introduction	2
1	Direct Coordination and Perceptualism	4
2	The Communicative Suggestion and Griceanism	9
3	Communication and Social Motivation	16
3.1	The Necessity Worry for Griceanism	18
3.2	Direct Coordination and the Concept Acquisition Problem	23
4	Conclusion	27

*This paper is accepted at *Inquiry*. I am grateful to Nate Charlow, Brendan de Kenessey, Imogen Dickie, Evan Westra, and an anonymous referee for comments on earlier versions of this paper. I am also grateful to John Campbell, Johannes Roessler, and Hong Yu Wong for comments at an APA Pacific Symposium in 2021, as well as participants at the MK40 workshop.

o Introduction

Joint attention occurs when two (or more) individuals attend together to some object (I focus on the two person case, as is standard). It is an early form of human joint engagement, and has been studied as such by a tradition of developmental and comparative psychology with roots in the pioneering work of Jerome Bruner and colleagues in the 1970s (Scaife & Bruner 1975, Bruner 1975a,b, 1977, 1983). From the psychological point of view, it provides a crucial milieu for cognitive development, especially that of social cognition; it is an interaction that is simple enough for infants to engage in, yet rich enough to foster cognitive growth.

More recently, philosophers have turned their attention to the phenomenon (two seminal works are Campbell 2002 and Peacocke 2005). For philosophers as well, theoretical interest in joint attention lies in its being at once sophisticated and simple, but in slightly different ways: the sophisticated element is its epistemological profile; the simple, its predominantly perceptual nature. The epistemic component of joint attention has been called its “openness”: the activity of the participants is mutually manifest or transparent—*out in the open*—between them. So, where psychological interest has been in its role in development, the philosophical puzzle has been to account for its openness as at least largely perceptual. But these foci are clearly related; the broad upshot of both the philosophical and psychological perspectives is that engagement in joint attention provides us with a basic grip on other minds, where “basic” is meant here roughly in the sense that sophisticated conceptual resources are not involved.

In addition, an exciting and very recent trend in both philosophical and psychological work on the topic consists of developing the suggestion that joint attention holds partially in virtue of communication (Carpenter & Liebal 2011, Eilan 2015, Siposova & Carpenter 2019, León 2021). Let me intuitively motivate this suggestion, as the cited authors do, with some contrived cases, which will also serve to pre-theoretically introduce our phenomenon of interest. After doing so, I explain what makes the communicative suggestion so interesting.

Suppose you are walking in the park and see a cute dog. The owner stands a short distance away. At first, you and the owner attend to the dog individually without noticing each other. Here are two ways things might unfold.

- i. You cast a furtive glance to the owner; you want to check if she is looking away so that you can get a chance to pet the dog without the trouble of asking for permission. She in turn notices your presence and subtly shifts her eyes towards you. Furthermore, suppose that each of you notice that the other is spying on the other. But, to avoid any awkward interaction, you covertly monitor her attention to you, and she does the same. Now, you are both aware of each other’s attention to each other and the dog. And potentially—through increasingly sneaky looks—you are each aware of each other’s awareness of that awareness, and so on. So here there may be arbitrarily high-level mutual awareness of your attention to each other and the dog; nonetheless, you are not attending *together* to the dog.

- ii. Upon seeing the dog and being overcome by its adorableness, you look to the owner in expression of your fondness for it. You smile and wink in that perfectly ambiguous way. Your eyes meet and perhaps you are happy to find her express a similar feeling, though with a slight tinge of weariness—presumably over the dog’s unbounded energy. Or perhaps you are surprised to find her totally fed up with the dog’s naughty behaviour. In any case, you look back at the dog in consideration of her reaction, expecting the interaction to continue.

Joint attention, along with its characteristic openness, is present in ii, but not i. And the contrast between these cases intuitively motivates the suggestion that communication plays a crucial role in sustaining joint attention.

Given its intuitive motivation, the communicative suggestion challenges a widespread approach in philosophy and beyond that characterizes public information in terms of common knowledge. Common knowledge holds when *A* and *B* know that such-and-such, they each know that they each know that such-and-such, and so on (forever).¹ In case i, however, there may be common knowledge between you and the dog owner about your attention to the other and the dog. Yet, again, there is no joint attention. Hence, such common knowledge is not sufficient for the kind of publicity that holds in joint attention. As shown by ii, it is really communication that brings about the openness of the interaction. Moreover, there seems to be no good reason to rule out, ahead of inquiry, the possibility of communication doing so without requiring common knowledge. Thus, the communicative suggestion promises a fresh account—*independent of common knowledge*—of least some forms of publicity.²

The communicative suggestion also illuminates a way of making progress on the challenge of accounting for joint attention’s basicness. For—and this is the guiding insight of this paper—the subtle nature of joint attention might be captured by a suitably subtle notion of communication. But there is a significant amount of ground clearing that must be done in order to develop the communicative suggestion in such a way as to help explain joint attention’s basicness, especially since the common knowledge tradition has its talons in the orthodox Gricean account of communication.

So here’s the plan. In §1, I refine the considerations mentioned above about joint attention’s basicness, and lay out Campbell’s (2002) perceptualist account of joint attention, according to which joint attention is a purely perceptual affair consisting of experiences and associated perceptual mechanisms. The perceptualist account, as I explain, is naturally motivated by the considerations of basicness. But, in §2, I present Eilan’s (2015) version of the communicative suggestion, which is specifically an objection to Campbell’s perceptualism. I then begin developing the communicative suggestion by looking to Griceanism. I propose a version of Gricean communication that does not rely upon anything like common knowledge. The final §3, however, begins by raising

¹The notion of common knowledge goes back to Lewis (1969) and Schiffer (1972). For prominent uses of it beyond philosophy, see Fagin et al. (1995) and Clark (1996).

²For recent work challenging the common knowledge orthodoxy, some of which I return to as the paper progresses, see Jankovic (2014), Tenenbaum (2015), Lederman (2017, 2018), and Harris (2019).

a couple of issues for my version of Griceanism, and then draws together the considerations of basicness and the role of communication from the previous two sections in order to meet those issues. Crucially, I do not argue that my version of Griceanism meets those issues; instead, I take them to reveal that there is a related, but simpler form of communication, driven by a deep-seated social need, present in early episodes of joint attention. In conclusion, I discuss an upshot that the present approach to joint attention has for joint action and publicity in general.

1 Direct Coordination and Perceptualism

The central line of argument for a perceptualist account of joint attention is that any alternative would appeal to beliefs or knowledge of the participants, yet such an appeal would overintellectualize the activity and not respect the primitive way in which joint attention provides access to other minds. This core motivation is emphasized in the philosophical approaches of Campbell (2002, 2005, 2011, 2018) and Peacocke (2005),³ and here it is expressed by psychologist Tomasello (2009).

As the children play, they monitor the adult and her attention, and the adult monitors the child and the child's attention. No one is certain how best to characterize this potentially infinite recursion of monitoring, but it seems to be part of the infants' experience—at least in some nascent form—from before the first birthday. (69)

Here Tomasello claims that joint attention involves a perceptual analogue of common knowledge, given presumably that bona fide common knowledge is too cognitively demanding for infants. Now, following a recent discussion by León (2021), I refine this line of argument by appeal to some specific results from the study of the development of social cognition, or theory of mind: our broad collection of cognitive skills for considering the mental lives of others.

An important notion here and for the rest of the paper is that of a basic episode of joint attention. Here is an empirical claim about such basic episodes.

Primacy Research in the development of theory of mind has shown that children begin engaging in joint attention as early as 9–12 months of age, and it is partly in virtue of engaging in joint attention that they develop competencies in more complex cognitive skills, including some involved in theory of mind.

For overviews of empirical support, see Carpenter et al. (1998) and Moll & Meltzoff (2011).

One important aspect of theory of mind is so-called “perspective-taking”. This concerns primarily our abilities to engage with others' visual perspectives (Flavell 1977, 1992). It is plausible

³The view that Peacocke (2005) develops is ultimately non-perceptualist, but in a subtle way, since he appeals to a category of mental states that he labels those of “awareness”. This category includes experiences, but also mental states that are more sophisticated than experiences, yet not as sophisticated as beliefs. I do not discuss the subtleties of this view, given space constraints. Campbell's perceptualist view is the focus of this section.

that joint attention always involves some sort of perspective-taking, and it might be tempting to think that they way it does is by a participant reflecting upon theirs and the other's perspective and recognizing that they are of the same object. The following empirical result, however, shows that basic episodes of joint attention do not involve that form of perspective-taking.

Confronting Perspectives Research in the development of theory of mind has shown that not until 4—5 years of age are children able to solve certain tasks that require them to CONFRONT PERSPECTIVES: simultaneously compare differing perspectives on a single object.

In support of this claim, a central study is from Moll et al. (2013).⁴ They presented children with a blue object, and a yellow colour filter that, when looked through, makes the object look green. Children were given experience with the object and filter to discover its effect. Then, in a situation where there is an adult looking through the colour filter at the object, and where the child sees the object without a filter, the child is asked at once both how the object appears to them and how it appears to the adult. The child responds by either answering verbally or pointing to a colour patch. The experimenters found that 3 year olds respond that it appears blue to both them and the adult, but 4–5 year olds answer correctly that the object appears blue to them but green to the adult.

The empirical conclusion that arises from Primacy and Confronting Perspectives is that children are able to engage in joint attention before they are able to confront perspectives. This conclusion suggests that it is partly in virtue of their engagement in basic episodes of joint attention that they develop sophisticated perspective-taking abilities. So, if joint attention always involves some form of perspective-taking, then this suggestion may be unpacked further as the claim that basic episodes of joint attention provide a primitive way of *sharing* a perspective with another that precedes and aids in the development of the ability to explicitly consider the other's perspective and compare it with one's own. This suggestion—that in basic episodes of joint attention there is a primitive sharing of perspectives—has informed recent psychological work on the topic (Moll & Meltzoff 2011, Tomasello 2018, 2019), but now I wish to draw out a philosophical upshot concerning justification, which provides preliminary motivation for the perceptualist account.

The distinction between sharing and confronting perspectives should bring to mind Frege's puzzle and his resulting distinction between names that share a sense and ones that differ in sense yet are nonetheless co-referential (Frege 1960). Upon understanding the sentence "Cicero is Cicero", it strikes one as trivially true because there is no sensible question of whether the two occurrences of "Cicero" stand for the same individual. According to Frege, this is so because the name "Cicero" has a single sense (which all of its occurrences share). But one may rationally wonder whether "Cicero is Tully" is true, since "Cicero" and "Tully" differ in sense, despite the fact that they pick out the same object.

⁴For a recent overview of empirical support for both Primacy and Confronting Perspectives, see Tomasello (2018, 2019). As Tomasello outlines, other studies with results complementary to Moll et al.'s (2013) are Flavell et al. 1981, Doherty & Perner 1998 and Rakoczy et al. 2015.

Instead of appeal to sense, a simpler and more modern way of addressing the phenomenon that Frege hit upon is in terms of coordination. Schematically, Frege’s puzzle was originally about how sentences of the form $\ulcorner a = a \urcorner$ and $\ulcorner a = b \urcorner$ could differ in cognitive value, where $\ulcorner a = b \urcorner$ is true. This puzzle can be generalized to be about how $\ulcorner \phi[a] \urcorner$ and $\ulcorner \phi[b] \urcorner$ can differ in cognitive value when $\ulcorner a = b \urcorner$ is true. For instance, someone might understand both “Cicero was a roman orator” and “Tully was a roman orator”, yet rationally accept one and not the other. As already glossed, in order to explain such differences in cognitive value, Frege posited the notion of sense. But what is sense? Talk of coordination allows one to capture rational relations between representations without reifying senses (Fine 2007). Instead of two expressions sharing a sense, one can say they are directly coordinated. Two co-referential singular terms are directly coordinated when understanding both guarantees justification for taking them to be the co-referential; two co-referential singular terms are indirectly coordinated when understanding both does not guarantee justification for taking them to be co-referential.

Setting aside relations between linguistic expressions, more relevant to present purposes are phenomena amenable to coordination at the level of thought: the mental states of a single subject. An old insight is that there is direct coordination between beliefs formed when attending to an object over time (Evans 1982, 1985, Campbell 1987). Suppose that you are watching your cat slowly walk around the room: you first notice how it raises its shoulders, as if ready to pounce on an imaginary target; a moment later you see it sniffing one of your plants, and hope it does not poison itself again. In a standard episode of solo attention such as this, where different properties of the object are noticed over time, there is immediate justification for grouping together all of the resulting beliefs as beliefs about the same object. That is, within such a procession of experiences, one does not have to constantly reflect and determine on that basis that an object seen at one moment is the same as one seen a moment later. You do not have to think, for instance, *this thing approaching the plant is the same thing as was prowling a moment ago, since I was watching it the entire time as it followed a natural trajectory for an animate object, and I may assume that there is no evil demon that somehow swapped in a similar looking cat along its path*. Some have suggested that the immediate justification one has holds simply in virtue of the phenomenal character of the experiences themselves (Campbell 2002, Smithies 2011, 2019). The PHENOMENAL CHARACTER of a mental state is what it is like for the subject to undergo it. Furthermore, perceptual phenomenology is TRANSPARENT: in the case of an experience, what it is like to have it is how things appear to the subject of the experience (Strawson 1988). The suggestion is thus that in sequences of experiences, modulated by attention, it is simply how things appear that there is an external world of enduring objects.

But the notions of coordination most relevant to present purposes are not intrapersonal, but interpersonal: they hold between one subject’s and another’s experience, in the context of a joint activity, when the subject has justification for taking hers and the other’s experience to be of the same object. DIRECT INTERPERSONAL COORDINATION in the case of joint attention involves a participant having immediate justification for taking hers and another’s experience to be of the same

object. As above, justification is IMMEDIATE when it does not rest upon conceptual reflection. So there is INDIRECT INTERPERSONAL COORDINATION when the justification rests upon conceptual reflection. These interpersonal notions are the most relevant to what follows, yet the discussion of the intrapersonal notions in the paragraph above was not merely an aside: Campbell's perceptualist account of joint attention, soon to be presented, draws a tight analogy between the two (and I return to critically discussing this analogy in the final section, when my own view is on the table).

With these preliminaries in place, I establish a constraint on basic episodes of joint attention.

Premise 1 Any episode of joint attention either involves direct or indirect interpersonal coordination between experiences.

Premise 2 If an episode of joint attention involves indirect coordination, then it involves deployment of the ability to confront perspectives.

Premise 3 In basic episodes, joint attention does not involve deployment of the ability to confront perspectives.

Conclusion In basic episodes, joint attention involves direct interpersonal coordination between experiences. [From P₁–P₃]

I justify the premises in turn. Premise 1 reflects an assumption that our engagement in joint attention involves a normative component. This assumption seems plausible if one considers what occurs in ii from the introduction, where there clearly is joint attention. However exactly the interaction goes, what you and the dog owner are doing—looking back and forth between each other and the dog, commenting on it in some way—makes it apparent that you are focusing together on the same dog. So I assume in general that in any episode of joint attention there is justification available to the subjects for taking their perceptual perspectives on the relevant object to be co-referential. And if this justification is not direct, then it must be indirect. Premise 2 holds because, in order for an individual to have justification for taking theirs and another's experiences to be co-referential, which is based on conceptual reflection upon two experiences, the individual must deploy the ability to simultaneously compare two perceptual perspectives. Premise 3 is supported by the empirical claims in *Primacy and Confronting Perspectives*: children are able to engage in joint attention before they have the ability to confront perspectives, and the development of the ability to confront perspectives unfolds partly in virtue of engagement in joint attention. So in these most basic cases of joint attention the ability to confront perspectives is not involved, and I label this conclusion the “direct coordination constraint” on basic episodes of joint attention.

The direct coordination constraint refines talk of joint attention's basicness, and thus also refines the central motivation for perceptualism about joint attention. For, assuming that experiences are personal-level but nonconceptual mental states,⁵ we can perhaps see our way to an

⁵The assumption that experiences are nonconceptual is controversial. For early versions of the view, see Dretske

account on which a proper configuration of experiences grounds the relevant epistemic facts of direct coordination.⁶ But, before presenting Campbell's perceptualism, let me pause to explain the background distinction between conceptual and nonconceptual mental states, as well as the related distinction between the personal- and subpersonal-level.

Beliefs and intentions are the paradigm conceptual mental states. For instance, I must have the concept of a frittata in order to believe that frittatas are tasty, or to intend to make a frittata for dinner. Experiences, I assume, are nonconceptual mental states. For instance, if you set a frittata down in front of me, I do not need the concept of one in order to see it. Furthermore, pre-experiential perceptual processing involves nonconceptual mental states: edge-detection involves calculations employing multi-variable calculus (Marr 1982), which my visual system performs with ease, despite that I struggled in first-year calculus. In general, following Peacocke (1992), a *CONCEPTUAL* mental state is one that, in order to be in it, the subject must possess the concepts required for the specification of its content.

Plausibly, any conceptual mental state is attributable to the person. For instance, it is I believing this frittata in front of me is tasty (so this clearly conceptual state is also personal-level); or, considering the contrapositive, it is not I calculating where the edges of it are on the basis of the retinal arrays of light intensities, since I do not know how to do those calculations (so this clearly subpersonal state is also nonconceptual). Yet there are nonconceptual mental states that are attributable to the person, the paradigm of which are experiences. For instance, it is I seeing the frittata. In general, a *PERSONAL-LEVEL* mental is one that is attributable to the person who is in it; whereas, a *SUBPERSONAL* one is instead attributable to one of their cognitive or physiological subsystems (Dennett 1969). Note that a crucial point later on in this paper is that there are nonconceptual personal-level mental states other than experiences.

Returning now to the main thread of the discussion, Campbell's perceptualist proposal is that joint attention is a three-place experiential relation between two individuals, who are present to each other as co-attenders, and an object. The relation is experiential in that, when two individuals and an object stand in it, the individuals are in experiential states. Furthermore, the object and the other's presence as co-attender are parts of each individual's experience. These experiential states are sustained by subpersonal perceptual mechanisms that track the object and the other's gaze; however, while the holding of the relation is sustained by subpersonal perceptual mechanisms—which may be deployed whether or not there is an object present or another individual present as co-attender—it is a relation that only holds if an object and co-attending individuals are in fact present.

(1981) and Evans (1982). More thorough motivation for and clarifications of that view are given by Peacocke (1992) and Heck (2007). A prominent defense of the view that experiences are conceptual is by McDowell (1994), but see Heck (2000) for a response.

⁶One immediate worry is that the direct coordination constraint rules out a perceptualist account of joint attention, since the constraint entails that in a basic episode participants must have the belief that theirs and the other's experiences are of the same object. But that entailment does not hold, since the relevant justification is propositional, and not doxastic. The justificandum of direct coordination is the content of a belief, but having justification for a belief-content does not require that one in fact form the corresponding belief.

This proposal faces an immediate problem: it is underexplanatory (Peacocke 2005, Eilan 2015, Battich & Geurts forthcoming). The other must enter one's experience *as a co-attender*, and not merely as, say, a parallel attender. But Campbell does not explain what it means for another to be present in one's experience in that way.

A natural way of elaborating Campbell's view in response to this initial problem, however, was anticipated by the discussion of intrapersonal coordination above. On Campbell's account of an individual's lone experience of an object, the role of subpersonal perceptual information-processing is in maintaining the subject's view of the object. A common way of understanding the role of this visual information-processing is in line with representationalism (see e.g. Harman 1990), on which the processing is glossed as the construction of a representation, which then exhausts the phenomenal character of the subject's experience. In contrast to the representationalist metaphor, Campbell suggests that the processing's role is simply that of making objects and their properties visible to the subject. Imagine looking through a pane of glass. The clearer the glass, the clearer the view; and there is no representation on the glass of what is seen through it. So the alternative metaphor is that the role of subpersonal perceptual processing is simply in maintaining the pane's transparency.⁷ Thus, by analogy with the lone case, it may be that in joint attention the subpersonal mechanisms of gaze-following make the other attender's activity transparent. Hence, the natural elaboration is as follows: when there is an object and another individual present, the deployment of the subpersonal perceptual mechanisms of gaze-following are sufficient for each individual's experience to contain the other as co-attender.

So, on the natural elaboration, there is a tight connection between solo and joint attention. In the solo case, subpersonal perceptual mechanisms, such as those bound up with edge-detection, play their role in maintaining a subject's view of an object in the world. In the joint case, additional mechanisms swing into play under the hood to track the other's attentional state. These additional mechanisms may be in a computational sense more sophisticated than those in the solo case, yet they are crucially of the same kind as pre-experiential subpersonal processing. It is for this reason that Campbell's account is perceptualist. But, it is also for this reason that the account is open to another, more prescient, objection. I begin the next section by presenting that objection, which motivates the communicative conception of joint attention.

2 The Communicative Suggestion and Griceanism

Eilan (2015) argues that even Campbell's elaborated proposal fails to distinguish another's presence as a co-attender, as in truly joint attention, from presence as a mutual covert attender. She supplies a pair of cases, one of which is a counterexample to the elaborated proposal. The cases i and ii I

⁷In the terms introduced by Chomsky (1995), this way of drawing the contrast between representationalism and direct realism presupposes an "ersatz" view of some mental representation, since it appeals to metaphorical glosses on the role of subpersonal processing. Compare Egan (2014, forthcoming). Though, following Campbell, I am careful to restrict ersatzism to subpersonal mental states.

gave in the introduction are similarly structured, and are partially based on her discussion.⁸

Suppose you are at a mandatory meeting in which a university administrator is droning on about some new policy. Now contrast the following two continuations. In the first, you look up and make eye contact with your colleague across the table and look back together at the presenter, perhaps exchanging looks of boredom. In the second,

you become aware, out of the corner of your eye, that a colleague is watching you. As you become aware of his observation of you, you start employing something psychologists call ‘covert attention’ with respect to him. And as you begin to do this, he, in turn, becomes aware that you are attending to him, and thus you continue for a while, each of you dividing your attention between the speaker and each other. (Eilan 2015: 6–7)

It is plausible that in both the first case, where there is joint attention, and in the second, where there is not, the same subpersonal mechanisms of gaze-following swing into play. As emphasized, these mechanisms are of the same kind as those undergirding lone attention. Hence, it seems plausible that they could be deployed in both types of case. But then the natural way of extending Campbell’s account does not provide a sufficient account of joint attention. The account predicts that another may be present as *co-attender* without there being joint attention.

Intuitively, what is missing in Eilan’s board meeting case 2, and present in 1, is communication. So let me now present a general framework that captures the communicative conception of joint attention thereby motivated. After discussing the framework, I begin the task—which is the concern of the rest of the paper—of filling it out.⁹

So here is my attempt at systematizing the general conception of joint attention that is motivated by Eilan’s objection to perceptualism.

The Communicative Conception of Joint Attention Individuals I and I' are jointly attending to an object o if and only if they are undergoing respective experiences E_I and $E_{I'}$ that are sustained by:

⁸As cited in the introduction, Carpenter and colleagues were the first to provide contrasts of this kind, though they do not specifically target Campbell’s perceptualism.

⁹Note that Seemann (2010) elaborates Campbell’s perceptualist proposal in order to meet the initial objection of underexplanatoriness, and the elaboration might also seem to be able to meet Eilan’s more advanced objection. Seemann, drawing on the psychological work of Hobson (2002, 2005), holds that the sharing of feelings about the object is part of joint attention (some other discussions that draw on this material are in Roessler 2005 and Eilan 2005). Crucially, however, the sharing is automatic and perceptual: one perceives another’s bodily expression of a feeling towards the object, which is then automatically replicated in oneself. So the sharing of feeling to which Seemann appeals is not a form of communication, since communication is an action. As elaborated much later in this paper, an action is something we do—not something that happens to us—and hence involves personal-level motivational states; whereas, Seemann (and Hobson) seem to be discussing a subpersonal and automatic process. Perhaps appealing to the mechanisms of the automatic echoing of affect makes it plausible that a different kind of subpersonal perceptual processing is being deployed between cases like board meeting 1 and 2. However, it seems plausible that the automatic sharing of affect could occur in cases like board meeting 2, where there is merely symmetrical covert attention: you may, for instance, feel the automatic pull of the other’s strange excitement about the presentation, despite the fact that you have only noticed it in a covert manner.

1. attention to the object *o*,
2. monitoring of the other's attention, and
3. communication between *I* and *I'* about *o* (*the communicative condition*).

The first two conditions follow Campbell in capturing the perceptual nature of joint attention. When individuals are jointly attending to some object, their experiences are sensitive to the object and each other's experiences. Both of these conditions can hold in virtue of subpersonal perceptual processing: our visual systems can track an object and another's gaze in a fast and automatic way. In an individual's lone attention, subpersonal perceptual processes play a causal information-processing role, and in virtue of doing so sustain a view of the external world for the subject via experience's phenomenal character. Analogously, as captured in conditions 1 and 2, in an episode of joint attention there are perceptual mechanisms of attention and gaze-following that help generate the phenomenal character of each participant's experience. But, crucially, the final condition represents that communication also plays a role in sustaining the experiences.

When stated at this broad level, the communicative conception does not necessitate the deep challenges to the common knowledge tradition mentioned in the introduction. In order to substantiate those upshots—how communication can bring about a kind of openness that is independent of common knowledge, how capturing that requires revising the traditional Gricean view of communication—the communicative component of the analysis must be elaborated. But, first, let me address some initial reservations one might have about the general framework.

According to some, certain cases of demonstrative communication hold partly in virtue of joint attention (Campbell 2002, Dickie & Rattan 2010, Seemann 2019). These cases are part of ordinary face-to-face conversation, and such that the referent of the demonstrative is perceptually available to the interlocutors. For instance, you point at a parked Ferrari and utter "That car is my favourite"; intuitively, your friend—though perhaps annoyed by your obsession with sports cars—understands your uttered demonstrative partly in virtue of you both jointly attending to the car. But the communicative conception flips this order of explanation. Suppose you are standing beside the same friend looking out a window. A loud crash reverberates from the street below. Both of you may look and see the passing truck responsible, and even notice that the other is doing so. Intuitively, however, there is no joint attention until there is communication between the two of you: perhaps a vocal expression of surprise, or simple raised eyebrows.

There is not the space here to provide a satisfying discussion of the relative priority of joint attention and demonstrative communication. The present point is simply that the communicative conception's verdict has intuitive support. In addition, the intuitive considerations at the end of the previous paragraph show that this conception is not in tension with the distinction between "top-down" and "bottom-up" joint attention (Carpenter & Liebal 2011: 170–171). In a top-down case, one participant draws the other's attention to the object, and such cases are naturally covered by the communicative conception of joint attention. In a bottom-up cases, however, an object makes itself salient by, for instance, making a loud noise. Such cases might seem to pose a problem for

the communicative conception; however, as illustrated above, despite the fact that an object might draw attention to itself, there is intuitively no joint attention to that object until the participants communicate about it in some manner.

I turn now to unpacking the communicative condition, which places an important constraint on theories of communication. Recall the contrasts given already as motivation for it: the one between i and ii in the introduction and the one in Eilan's board meeting. Each of these contrasts includes a case where the perceptual components 1 and 2 of the communicative conception are satisfied, yet there is no joint attention. Hence, if it is the addition of communication that brings about joint attention, then successful communication must itself be sufficient to help bring about the jointness of an activity. A simple way in which communication might be sufficient in this regard is if communication is itself a joint activity, and I assume that this simple way in which communication brings about jointness is correct. Hence, the constraint that the communicative conception of joint attention places on theories of communication is that communication must be represented as a joint activity. After all, the claim that communication is something we do together has intuitive support independently of the present considerations about joint attention; the communicative conception of joint attention simply brings into focus the joint nature of communication.

Let me begin the search for an account of communication that respects its status as a joint activity by reflecting upon joint activity in general. According to the influential view of joint action of Bratman (1993, 2014), common knowledge of interlocked intentions is sufficient. On this view, for instance, we are walking together if the following holds: we both intend that we walk and that our resulting sub-intentions—to take one step, then to take a second—should influence one another, and all of this is common knowledge. But here is a case from Tenenbaum (2015) that forcefully illustrates how Bratman's account falters even for the pedestrian case of walking together.

Let us suppose Larry has an inordinate amount of money coming to him. But the bank will give Larry the money only if he shows up at the bank with Mary so that a bank teller can confirm that Mary is alive. Unfortunately, as Larry knows, Mary would prefer him to be penniless, so she's not willing to go to the bank with Larry simply to help him. However, as luck would have it, Larry finds out that Mary is in the perfectly symmetrical predicament. As one would expect, Larry feels the same spitefulness toward Mary. Their mutual hatred runs deep but not as deep as their self-love; each would rather have the money rather than let the other suffer in poverty. Unfortunately, more powerful than their mutual hatred or their inflated self-love is their sense of dignity. Neither would stoop to ask the other a favour or propose a truce or an agreement. They all know all this and are thus incapable of retrieving their respective fortunes. One day Larry is walking and he sees Mary; he immediately realizes that if he walks towards the bank, she'll follow him there in the hopes she'll

get the money. Mary is hit by the same thought. They immediately notice that both would have had the same thought, and that they both realize that the other would have had noticed that both had the same thought, etc. They walk towards the bank, each carefully monitoring that the other is going. (3385)

Here each of Larry and Mary intend that *we walk to the bank*, and that their steps influence one another's. Furthermore, there is common knowledge of this whole affair. Yet there is intuitively no joint action: they are not walking together to the bank.¹⁰

Tenenbaum suggests that what is missing is that Larry and Mary are not intending to do something that is essentially joint, i.e. an activity that cannot be done alone. They are not, for instance, intending to walk *together* to the bank. Furthermore, once the intentions to engage in some such joint activity are added, there is no need to insist upon a requirement of common knowledge. The present suggestion is thus that it is necessary and sufficient for two individuals to be walking together that each are walking *because* both intend that they walk together. Thus, the general proposal—which I accept in what follows, though with one subtle modification that I will mention in §3.1—is that two individuals are doing *X* together just in case they each doing *X* because both intend to do *X* together. It is crucial that the relevant sense of “because” here is not a merely causal one; rather, it must be the sense in which an agent's action is explained as an intentional action. There need not, however, be a requirement of common knowledge added regarding the intentions to act together. Tenenbaum's suggestion thus provides an account of joint action according to which it is orthogonal to common knowledge, and such an account is supported by intuition: the case above of Larry and Mary suggests that common knowledge is not sufficient; there are also intuitive cases of joint action where common knowledge is not possible.¹¹

One might have an at least *prima facie* worry that Tenenbaum's suggestion about the nature of joint action is circular—as Tenenbaum himself does, following Bratman (2014)—for the holding of a joint activity is to be explained in part in virtue of an intention with a content whose specification involves appeal to a concept of that joint activity. But there is nothing in principle incoherent or underexplanatory about this form of circularity: as Fine (2012) observes in a related discussion, perhaps all there is to being cool is to be taken to be so. So, I wish to connect Tenenbaum's suggestion with some central aspects of the Gricean approach to communication in order to give an account of communication that secures communication's status as a joint activity. I return to the subtle circularity worry after developing a refined Gricean view of communication and contrasting it with the standard version of Griceanism.

For Grice ([1957] 1989), the speaker's part of an episode of communication includes an intention to bring about a certain change in the addressee's mental life. But this intention—called the “informative” intention—is not the whole story: a parent might physically discipline their child

¹⁰A related point has been emphasized by Gilbert (1990, 2013), though she focuses on the claim that common knowledge of interlocked individual intentions is not sufficient for the socio-normative commitments intuitively present in joint action. Following Tenenbaum, I do not take the upshot of the case given here to be primarily about joint commitments.

¹¹For the latter type of case, see, for instance, Jankovic (2014: 505).

and thereby make the child believe that they acted wrongly. But that morally dubious use of force is not communication. In addition, the speaker must have what is called a “communicative” intention: as a first gloss, the speaker must intend that her informative intention be recognized by the addressee and for that recognition to serve as at least a part of the reason for the addressee’s bringing about of the mental change. But that first gloss is not the whole story: a neuroscientist might have a machine that can manipulate brain waves via electrical shocks in order to bring about certain mental states. The neuroscientist might have an informative intention and use this device to bring about others’ recognition of it. But in such a case there is intuitively no communication, even if the neuroscientist intends that the unfortunate “addressee” bring about the intended mental change with the induced recognition as a reason for doing so.¹²

One might (correctly) suspect that the notion of recognition should rule out the neuroscientist case: the speaker, *inter alia*, intends that the addressee *recognize* their intention. But an account is needed of what that recognition involves. I propose to use the notion of acting together, taken from Tenenbaum, as the basis for the required account: in an episode of joint action, there is a joint goal, which is a joint activity represented in the intentions of the participants. I claim that this joint goal is that of mutual intelligibility: making sense of one another as fellow agents. So the speaker’s communicative intention involves the aim that she and the addressee work together in making sense of one another’s action, and that successful engagement in this activity generates the addressee’s recognition of the informative intention.

Here is the proposal stated in systematic fashion.

Refined Gricean Communication There is communication_{fg} (in the joint-activity-based Gricean sense) between speaker *S* and addressee *A* just in case *S* makes an utterance where:

1. *S* intends that a particular reaction *r* is brought about in *A* (*the informative intention*),
2. *S* intends to engage with *A* **in the joint activity of making sense of one another’s actions**, where successful engagement in this activity generates *A*’s recognition of the informative intention (*the primary communicative intention*)
3. *S* intends that this recognition serve as at least a part of *A*’s reason for producing that effect (*the secondary communicative intention*), and
4. the primary communicative intention is fulfilled (*the uptake condition*).

It is crucial to note that, as indicated by the subscript of *fg* on “communication” in the analysis above, I have only provided necessary and sufficient conditions for a specific type of communication: the joint-activity-based Gricean one. I am not claiming that there are no other forms of communication. Nonetheless, I do think that the constraint that communication is a joint activity, which was brought to light by the communicative conception of joint attention, shows that

¹²This type of case, which appeals to a neuroscientist that can bring about mental states with shocks to the brain, is from Schiffer (1972).

the standard version of Gricean communication is incorrect. For that reason, which I am about to elaborate, I take my proposal to be a superior successor to standard Griceanism.

In order to compare my proposal with standard Griceanism, let me first say a few words about the concept of the joint activity of making sense of one another, which I have claimed is crucial for specifying the content of the speaker's primary communicative intention. The paradigm instance of making sense of another's actions is the attribution of conceptual mental states such as beliefs and intentions. And that paradigm is exactly the intended result of the speaker's primary communicative intention, which is that the addressee attribute an intention to her. But simpler forms of making sense of one another should be possible: one might use whatever cognitive resources one has in order to see how another is "like me" (Tomasello 1999: 70). And these simpler forms are relevant to how that overall result—the attribution of the speaker's intention—is brought about (in addition to the potential attribution of other intentions or beliefs in bringing about that result). The addressee should, according to the speaker's primary communicative intention, work together with the speaker in order to discover the informative intention, in such a way that indicates to the speaker that the addressee is a fellow agent. I return to this particular joint activity in the next section, as I will argue that the concept of it is gained from a deep-seated social motivation.

On the standard Gricean picture, the bold part of the primary communicative intention above is instead filled out by appeal to common knowledge: the speaker intends that her and the addressee come to have common knowledge of the informative intention. So communication is constituted by common knowledge of the speaker's attempt to influence the mental life of the addressee.¹³ In the present context, however, one should be suspicious about whether the appeal to common knowledge helps provide sufficient conditions for communication. Recall the cases given in support of the communicative view of joint attention—Eilan's board meeting example, case i from the introduction—there is common knowledge between the relevant individuals of their attention to each other and some object. Yet the cases illustrate that such common knowledge is consistent with a lack of genuine engagement between them. Thus, it should be possible for there to be common knowledge of a speaker's informative intention, yet no genuine joint activity. After all, as observed by Eilan (2020), Grice's original account of communication represents its phenomenon as a relatively impersonal affair: the speaker acts with certain intentions, which the hearer then recognizes. And, since common knowledge does not suffice for joint activity, adding common knowledge of the speaker's intention should not be expected to make the overall account sufficient. So it seems that the standard Gricean view simply pushes back a step the problem for common knowledge raised by the cases motivating the communicative suggestion

¹³This standard view is from Schiffer (1972). An important alternative version of it is given by Sperber & Wilson (1986), who appeal instead to mutual manifestness, which is a weaker notion than common knowledge. But I will not elaborate upon their alternative, since the difference between the two views does not matter for present purposes: the criticism below of the common knowledge orthodoxy I am about to present is that it does not provide sufficient conditions for communication, so appeal to a notion weaker than common knowledge will not address that criticism. This same point applies for the even weaker version of Griceanism in Neale 1992.

about joint attention.

As an alternative, I have proposed a refinement of Griceanism that is in line with various recent proposals that communication should be a genuinely joint endeavour (Jankovic 2014, Siposova & Carpenter 2019, Eilan 2020). Crucially, my proposal draws on Tenenbaum's suggestion that genuine joint action requires that the participants intend to engage in an activity that is essentially joint. In the case of communication, again, I propose that the intended activity is that of making sense of one another's action, where this activity is understood as something that must be done together.

In order for my alternative to ultimately prove a genuine successor to the common knowledge orthodoxy, it must be shown that it can also handle the main motivation for appeal to common knowledge in the primary communicative intention. The main motivation is the problem of sneaky intentions. The problem of sneaky intentions may be dealt with on my joint-activity-based Griceanism, but I do not do so here—though see this footnote¹⁴ for a brief discussion. Instead, I turn now to the final section, which I begin by raising two issues for my refined Gricean view of communication. The resolution of these issues then guides the development of my overall account of the openness of joint attention and its special role in our social lives.

3 Communication and Social Motivation

Recall the circularity issue for Tenenbaum's suggestion: something, X , is claimed to hold partly in virtue of individuals being in mental states such that the specification of the content of those mental states involves appeal to a concept of X . My joint-activity-based Griceanism also partakes in that form of circularity. I already suggested, drawing on an observation from Fine (2012), that there is nothing in principle incoherent or underexplanatory about this form of circularity.

¹⁴The problem of sneaky intentions was first raised by Strawson (1964). It was then crucially elaborated by Schiffer (1972), who provided cases of increasing complexity that suggest the addition of common knowledge to fill out the speaker's primary communicative intention. Grice ([1969] 1989), however, suggested that there is in fact something else going wrong in all of the cases put forward by Schiffer. Grice observes that in such cases the speaker intends that the addressee rationally transition from recognition of the informative intention to its fulfillment in a certain way, but at the same time intends that the addressee falsely believe that the speaker intends him, the addressee, to *not* make the transition in that precise way. So, in such a case, the speaker acts with an intention that P , but at the same time intends that the addressee think falsely that she, the speaker, has an intention that $\text{not-}P$. So, Grice suggested that, instead of appealing to common knowledge, his account of communication may simply add the requirement that the speaker does not have intentions of that form. But an immediate problem with that suggestion is that it seems ad hoc. So let me suggest that my joint-action-based Griceanism promises to provide a principled way of ruling out the kind of scheme that Grice identified is present in cases illustrating the problem of sneaky intentions. In general, it is plausible that one cannot act with the intention of working together with another, with the aim of making each other's actions mutually intelligible, while at the same time intending that the other form a false belief about the intentions behind one's action. Finally, see Jankovic (2014) for a closely allied version of Griceanism, which stresses how communication is a form of joint action, and in which it is argued in detail that such a version of Griceanism can deal with the problem of sneaky intentions. Though it is a subtle question whether I can adopt her solution, since she stresses Bratman's account of joint action, yet I go along with Tenenbaum's alternative. Overall, I do not wish to get into the details of how the present account fully solves the problem of sneaky intentions; instead, what I say in response to the problem here is intended as an indication that the present view is on the right track as a replacement of the common knowledge one.

Furthermore, it is not even clear why these authors—Fine, Bratman, and Tenenbaum—even take there to be a *prima facie* issue here.¹⁵ There does not seem to be anything wrong with saying, for instance, that someone’s action counts as one of intentionally shutting the car door partly in virtue of them intending to shut the car door. Analogously, the actions of two people may be said to amount to the joint activity of, say, walking together partly in virtue of each of them intending to engage in that joint activity.

There is, however, a crucial difference between the examples just given. In the case of shutting the car door, there is a familiar story as to how one may come to possess the concepts required for specifying the intention’s content. By perceptually interacting with the world, one can learn about car doors and what it is for something to shut. In contrast, the acquisition of a concept of a joint activity does not seem to fit the familiar mold of learning from perceptual experience, for if such a concept is one that is gained, then it is gained partly in virtue of engagement in joint action—not mere observation of external reality. But then there is a conflict with Tenenbaum’s suggestion that engagement in joint action holds partly in virtue of the deployment of a concept, already possessed, of an essentially joint action. It cannot be both that joint activity holds partially in virtue of the deployment of a concept, where that concept itself must be gained partially in virtue of engagement in joint activity; otherwise, either engagement in joint activity is impossible or the concept is not one that is gained.

So any account of joint action that takes Tenenbaum’s suggestion on board faces an issue concerning how to properly account for we gain the concept of the joint activity that the account appeals to. Again, the issue arises because it is plausible that the concept of any joint activity is one that is gained partly in virtue of joint engagement. I label this issue the “concept acquisition” problem. Here is how this problem applies in particular to my account of joint attention and communication from the previous section. I have claimed that episodes of joint attention hold partially in virtue of a form of communication that involves deployment of the concept of the joint activity of mutual intelligibility. But it is plausible that the concept of that joint activity is gained partially in virtue of engagement in the joint activity of achieving mutual intelligibility—in basic episodes of joint attention, for instance. By the end of this section, a solution to the concept acquisition problem will be given on behalf of my account.

The second issue I address in this section concerns the necessity of any form of Gricean communication for sustaining joint attention. Recall the Primacy claim from §1: Infants begin engaging in joint attention very early on. So one might worry about attributing Gricean intentions and their recognition to infants.¹⁶ So, if the communicative conception of joint attention is correct, then the relevant form of communication in basic episodes of joint attention is not the joint-activity-based Gricean one, on pain of overintellectualization. It is additionally important, in the present context, to isolate the form of communication that is present in such cases, since, as was

¹⁵Thanks to Imogen Dickie for discussion on this point.

¹⁶For more developed arguments against Gricean communication along these lines, see Roessler (2005) and Campbell (2018).

argued in 1, the main motivation for the perceptualist view of joint attention is based on the considerations of basicness, and in particular the direct coordination that must be present in early episodes of joint attention. So it is worth investigating whether the communicative conception can provide an alternative account of direct coordination.

The plan for the rest of this paper is to use the necessity worry as a springboard: given that there is some intuitive pull to the claim about overintellectualization, addressing it leads the way to a communicative account of direct coordination—which brings together the seemingly independent considerations from §1 concerning joint attention’s basicness with those in §2 motivating the communicative suggestion—and also provides an answer to the concept acquisition problem.

3.1 The Necessity Worry for Griceanism

The first step in meeting the necessity issue is in generalizing the Gricean conception of communication by allowing there to be a form of genuine communication based on personal-level motivational states other than intentions. Recall from §1 that intentions are *conceptual* personal-level motivational states. Following Dickie (2015), I label *nonconceptual* personal-level motivational states “needs”. Paradigm examples of needs are hunger and thirst. Animals and infants may be so driven without possessing the concept of, say, nutrition. Personal-level motivational states in general, whether needs or intentions, have the common feature that they are intimately related to genuine action: an action—as opposed to a reflex, or something that happens to us—is a movement guided by a personal-level motivational state (Frankfurt 1978). Given the existence of needs, it is a short step to recognizing a form of communication that is genuine action, yet not based on intentions.

So I propose that there is a form of communication driven by a social need. The content of this need is to make sense of each other’s actions: we are driven to mutually establish one another’s intelligibility. In specifying the content of this need, I am appealing to the same concept as was used to specify the communicative intention of my joint-activity-based Gricean account of communication (recall how on that view the speaker must have the informative intention to bring about a change to the addressee’s mental life, but also the communicative intention that the addressee recognize the informative intention *by virtue of the joint activity of making sense of another*, and for that recognition to be part of the addressee’s reason for bringing about the intended effect). Though, again, needs are nonconceptual, so one can be driven by a need without possessing the concepts required for its specification.

That there is some form of communication occurring before and within basic episodes of joint attention is empirically supported by what psychologists call the transition from “primary” to “secondary” intersubjectivity (Trevarthen 1979, Rochat & Striano 1999, Adamson & Russell 1999, Hobson 2002). Throughout the first year of life, infants engage in one-to-one emotional communication with others: primary intersubjectivity. There is a huge amount of evidence that as early as the first couple months infants are active and aware participants in emotionally laden social

exchanges. See this footnote¹⁷ for some empirical detail. Here is a psychologist's description of a 2-month-old (Tronick et al. 1978).

Baby is looking off to side where mother will come in. He sits completely quiet, back in his baby seat, face serious, cheeks droopy, mouth half open, corners down, but there is an expectant look in his eyes as if he were waiting. His face and hands reach out in the same direction. As his mother comes in, saying, "Hello" in a high-pitched but gentle voice, he follows her with his head and eyes as she approaches him. His body builds up with tension, his face and eyes open up with a real greeting which ends with a smile. His mouth opens wide and his whole body orients towards her. He subsides, mouths his tongue twice, his smile dies, and he looks down briefly, while she continues to talk in an increasingly eliciting voice. During this, his voice and face are still, but all parts of his body point toward her. After he looks down, she reaches for and begins to move his hips and legs in a gentle, containing movement. He looks up again, smiles widely, narrows his eyes, brings one hand up to his mouth, grunting, vocalizing, and begins to cycle his arms and legs out toward her. (5)

In addition to one-to-one affective communication—the primary intersubjectivity illustrated in the passage just given—throughout the first year of life infants separately engage in gaze-following and lone attention to objects. Then, around the first birthday, all of these activities come together and infants begin sharing with others their emotive reactions to objects: the dawning of secondary subjectivity (Striano & Bertin 2005). Here is a psychologist's description of an early episode of secondary intersubjectivity (Stern 1985).

A nine-month-old girl becomes very excited about a toy and reaches for it. As she grabs it, she lets out an exuberant "aaah!" and looks at her mother. Her mother looks back, scrunches up her shoulders and performs a terrific shimmy with her upper body, like a go-go-dancer. The shimmy lasts only about as long as her daughter's "aaah!" but is equally excited, joyful, and intense. (140)

¹⁷In particular, there is the still-face effect, which was first established by Tronick et al. (1978) in studies involving infants between 2 and 20 weeks old and their mothers. The behaviour of the infants was compared in two kinds of interaction with their mothers: normal and still-faced. In the normal interaction, the mother played with the infant. Normal playing involves, for instance, miming of the infant's facial expressions, as well as intervening to coax it away from unhappy expressions, or to contain over-excited ones. In the still-face interaction, the mother sits in front of her infant and stares at it with a neutral face. In these cases, the infant attempts to in many ways to engage the mother, but after this is unsuccessful it withdraws and expresses wariness and helplessness. The still-face effect is the significant difference in infants' behaviours between these two different kinds of interactions. According to the meta-analysis of Mesman et al. (2009), similar results have been found in over 80 studies since Tronick et al. 1978. These studies vary with respect to the age and gender of the children, as well as whether the adult engaging with the child is a parent or stranger. Several studies have been done with non-Western children, and found the effect as well. In addition to the still-face effect, an important source of empirical support for the connection between the social motivation and early episodes of joint attention is that there are negative correlations between Autism Spectrum Disorder (ASD) and engagement in joint attention, among other aspects of social cognition, which supports the present proposal given the recent view that ASD is characterized by a deficit in social motivation (Chevallier et al. 2012, Nyström et al. 2019). For in-depth overview of developmental as well as evolutionary considerations in support of the social motivation's relevance for joint attention, and infant sociality more generally, see Tomasello (2014, 2019).

Described here is a basic episode of joint attention—the type appealed to at the beginning of §1. Finally, the active and aware engagement illustrated in the above passages is explained if the infants’ activity is genuine action driven by a motivational state that concerns others.

An action driven by the social motivation need not draw upon any intentions or other conceptual mental states. Compare action driven by hunger, which is another human need. In some cases, one forms beliefs and corresponding intentions about how best to satisfy one’s hunger. Imagine you are deliberating about which restaurant to visit. But sometimes one’s action is driven by hunger without any intervening conceptual mental states. Imagine your ravenous appetite has you pulled in, without reflection, by the smell of a familiar fast food chain. Similarly, I claim that the social motivation may guide one’s action in communicating without the formation of beliefs or intentions. In particular, one need not have the Gricean informative intention—a part of the Gricean account of communication from the previous section—to influence the addressee’s mental life in some way. In the simple form of communication I am here considering, the infant is driven by their need to make sense of and be made sense of by others in whatever way they can, which in early life involves the simple expression of emotion.

There are a number of recent proposals from both philosophers and psychologists similar to mine regarding a social motivation. So let me say a few critical words about these alternatives in order to sharpen my proposal about the social motivation’s content. First, Eilan (2015) speaks of a “desire for relatedness” or “basic human urge to connect” (12), which may be unpacked slightly as a need for social bonds. Yet even when there is a strong bond between two individuals—a child and caregiver, for instance; or simply two adult friends—there is continued drive to engage in joint activities. Thus, Campbell’s (2018) proposal that we have a need for “cooperative social interaction” seems superior (124), since instead of binary relatedness the need Campbell proposes has ternary content: we have a need to engage with others in joint activity.¹⁸ Campbell also proposes that the need might be specifically for joint attention, and a similar proposal is made by Dickie (2020), who claims that the need is to engage with others towards objects in the world in such a way as to mutually and rationally gain information about those objects. I take from Dickie and Campbell the insights that the social motivation has ternary content. But I hold that the ternary content does not appeal directly to things in the world about which we are trying to engage with others. Rather, the motivation is to engage in the joint activity of making sense of one another in light of our actions; so, there is more than simply the binary content of connecting with another, but the additional aspect does not directly concern independent objects. By not appealing to independent objects, as Dickie and Campbell do, the present account explains how the social need drives early one-to-one episodes of communication, and thus also how basic episodes of joint attention are related to those one-to-one interactions.

Dickie’s and Campbell’s proposals are closely allied with Tomasello et al.’s (2005) influential “Shared Intentionality Hypothesis”. This hypothesis is about what is unique about human cognition and its development, and it claims that it is our engagement in joint action. Accordingly,

¹⁸Thanks to John Campbell for discussion on this point.

the hypothesis includes the claim that we have a “species-unique motivation to share emotions, experience, and activities with other persons” (Tomasello et al. 2005: 675). As the hypothesis is elaborated by Tomasello (2009, 2019), communication is a slightly sophisticated form of joint activity, which comes after and partly in virtue of participation in joint attention. On my view, however, which takes seriously the communicative suggestion about joint attention, the most basic type of joint action is communication. Let me explain. Recall first the proposal about joint action from the previous section: two individuals are doing *X* together just in case each are doing *X* because both intend to do *X* together. In the beginning of the present subsection, the existence of personal-level motivational states other than intentions was revealed, so it is natural to broaden that account of joint action in the following way: two individuals are doing *X* together just in case each are doing *X* because both are in a personal-level motivational state with the content to do *X* together. On this broadening, it is not a requirement that both participants intend to perform the relevant activity together. Instead, since the only motivational states other than intentions are needs—again, intentions are the conceptual personal-level motivational states, and needs are the nonconceptual ones—it may be that at least one of the participants is acting under a need to do *X* together. Now, I am proposing that the social motivation is the only human need with the relevantly joint content. Thus, achieving mutual intelligibility is the only joint activity that can be engaged in without an intention with joint content. Furthermore, as outlined above, it is one that we engage in very early on in expressing emotions, and we do so without reliance upon any conceptual mental states. Finally, I propose that joint engagement in the activity of achieving mutual intelligibility itself is a form of communication. Thus, the only kind of joint activity that can be engaged in without any intentions—and which we do in fact engage in right out of the gate—is a form of communication.

So what I am claiming about the nature of basic episodes of joint attention is that they are simply the joint activity of achieving mutual intelligibility, but where there is the expression of emotive states concerning some object, which helps sustain mutually responsive experiences of attention to that object. Now, one issue for my view, which seems to favour the Shared Intentionality Hypothesis and its philosophical allies, can be put crudely with the following question: Why, in joint attention, does the object matter at all? There are many ways of making sense of one another: as already stressed, there is a kind of mutual intelligibility attained in virtue of the direct communication of affect—not about any object. So, if basic episodes of joint attention are driven by the need for mutual intelligibility, and infants throughout their first year of life are already engaging in direct communication of affect, then why is it that joint attention emerges?

The answer I propose draws on a general fact about action that is driven by needs. The fact is that not all ways of satisfying the content of a motivational state lead to a release of motivational pressure. Consider, for instance, the feeling of hunger and consequent satisfaction upon eating. In order for one’s feeling of hunger to be reduced, there are specific ways in which nutrients must be processed; if one were to place food directly into the stomach or inject calories, for instance, one would not feel less hungry. It might be suggested, in response, that specifying the content of

motivational states does involve specifying particular routes to fulfilment. Fully responding to this suggestion would take us too far afield from the present topic, since there are general issues about the content of motivational states that are relevant.¹⁹ Nonetheless, packing in, for instance, ways of getting nutrition into the content of the state of hunger seems immoderate. So I am hopeful that the intuition that the content of, say, hunger does not involve a way of gaining nutrients can be theoretically vindicated. But I leave such exploration to future work.

Applying this general insight to the present topic—the social motivation—I propose that, as an infant’s development proceeds, it becomes crucial for abating the motivational pressure of the social need that the activity be one of making sense of one another in relation to some perceptually available object. In these contexts, sharing attitudes about an object is crucial for releasing the relevant motivational pressure. One might further ask why the relevant motivational pressure has this feature. The answer to this further question presumably lies in the role joint attention plays in cognitive development. Within episodes of joint attention, there is cognitive scaffolding that parents and older peers provide, for which the presence of an object is more beneficial than what is available in simpler one-to-one engagement. So, from an evolutionary point of view, early social interaction is a means to the child getting into a position in which they have the independence to engage with the world. Nonetheless, from a developmental point of view, social interaction is the foreground, since it is a basic need for the joint activity making sense of others and being made sense of driving the child’s activity, and this need itself does not involve some further purpose.

The crucial point from the previous paragraph is that the causal role of motivational phenomenology explains the transition from primary to secondary intersubjectivity. In primary intersubjectivity, recall, there is direct one-to-one affective communication, which is one particular way of satiating the social need. Secondary intersubjectivity involves another particular way of satiating the social need, in which one communicates with another concerning some object. I claim that, as development proceeds throughout the first year, motivational phenomenology pushes one towards the latter way of establishing mutual intelligibility.

Now let me say more about motivational phenomenology, which is the feeling of pressure and release associated with needs. Philosophers have traditionally focused on perceptual phenomenology, but it should be uncontroversial that motivational phenomenology exists. There is something it is like to feel hungry, and sated. Similar phenomenology is associated with the social need: there is a way it is like to need to connect with others, and for this motivational pressure to be released when one engages with another human. Consider, for instance, the satisfaction one feels from doing even mundane activities with others. Suppose you are forced to spend months in social isolation (please forgive the outré philosophical thought experiment). During such a period, one “hungers” for social interaction. And, after such isolation, consider what it is like to finally sit and talk with a friend, or even a stranger. This particular phenomenology is generated from the

¹⁹The issue, for instance, of whether the content of a motivational state is an accuracy condition—how things will be—or a fulfillment condition—how things should be. See Dickie (2015: §3) and references therein for further discussion of this distinction. For empirical overviews of relevant empirical work on motivation, which support the empirical claims made earlier in this paragraph, see Toates (1986) and Berridge (2004).

motivational pressure and release of the social need.

Other defenders of the communicative conception of joint attention have stressed the importance of a similar feeling of connection had when communicatively interacting with others (Eilan 2015, 2018, Siposova & Carpenter 2019). So here is a point, as with many others above, where I am drawing together and refining recent psychological and philosophical work surrounding joint attention, with special attention to that of Eilan (2015, 2018, 2020). But, in the next subsection I show how the motivational phenomenology I have highlighted helps make progress with regard to the direct coordination constraint, from §1, and the concept acquisition problem raised at the beginning of the current section.

3.2 Direct Coordination and the Concept Acquisition Problem

Recall the direct coordination constraint: basic episodes of joint attention are interactions that make other minds transparent, which is to say that participants have immediate justification for taking theirs and the other's experience to be of the same object. That justification is immediate in the sense that it is not based upon conceptual reflection on ones and the other's perspectives on the object. This direct coordination constraint refined the introductory talk about joint attention's basicness, and thus is a desideratum for any account of joint attention.

I also explained in §1 how the direct coordination constraint naturally motivates Campbell's perceptualist account of joint attention, since transparency, according to perceptualism, is purely a matter of the perceptual phenomenology of experiences, which is nonconceptual. But Eilan's objection to perceptualism, given at the beginning of §2, casts doubt on the claim that the transparency of other minds in joint attention is purely a matter of perceptual phenomenology. Therefore, *if it is assumed that the only phenomenology relevant for direct coordination is perceptual*, then there must be a non-phenomenological component to direct coordination. Presumably, the non-phenomenological component would be the conceptual content of certain relevant mental states.

Eilan (2015) seems to make the assumption just emphasized, and is thus driven to claim that in joint attention one's experience has conceptual content. She characterizes this content as follows: *we (you and I) are sharing this experience.*²⁰ Crucially, this conceptual content involves a primitive second-personal way of thinking of the other, which is the concept that plays a crucial role in bringing it about that the interaction is genuinely joint. On her view, we are doing something together because we are thinking of each other under "primitive you-awareness" (6). But such an appeal to conceptual content brings along a worry of overintellectualization related to the necessity worry for Griceanism above. And, more importantly for present purposes, such an appeal means that Eilan's view cannot resolve the concept acquisition problem. If the second-person concept were one that is gained, it would plausibly be so in virtue of engagement in joint

²⁰See especially pages 15–16 of Eilan 2015. I explained above that experiences have nonconceptual content, but that claim is consistent with them having some conceptual content, for it may be that the way experiences contrast with, say, beliefs is that the latter do not have any nonconceptual content. Regardless, I am rejecting Eilan's view, and my view is consistent with the denial that experiences have any conceptual content.

activities such as joint attention. So, the sense in which Eilan calls the second-person concept “primitive” comes into view: it is not a concept that we gain.

But, as emerged in the previous subsection, there is relevant motivational phenomenology from the social need. Thus, we may follow Eilan, against Campbell, in accepting that perceptual phenomenology is not sufficient for the transparency of joint attention, yet at the same time reject the crucial assumption, which both authors seem to share, that the only phenomenology relevant for direct coordination is perceptual. Accordingly, as I show in the remainder of this subsection, we may provide a fully nonconceptual account of direct coordination as well as an answer to the concept acquisition problem.

Beginning with direct coordination, I propose that the justification had for taking the two experiences to be of the same object is given by the practical justification of one’s activity in jointly attending.²¹ Let me explain the background notion of practical justification. All actions have a certain baseline normative or rational status, since they are things we do and not mere worldly happenings. But, in addition, some actions have the further positive normative status of practical justification. Theoretical justification may be more familiar to many philosophers. Traditionally, theoretical justification is had by beliefs. For instance, in normal conditions, if one sees or is told that things are thus-and-so, one’s resulting belief that things are thus-and-so has the positive normative status of being justified. Contrast such a belief to one formed on a hunch. But it is also natural to talk of actions having similar normative status. Contrast the movements of a skilled archer in aiming at a bullseye with those of a novice. In general, we may say that an action has PRACTICAL JUSTIFICATION just in case it is appropriately related to its guiding motivational state. And, just as with theoretical justification, one might refine this appropriate relation along either externalist or internalist lines.

One aspect of practical justification is reliability. Consider how the skilled archer’s actions reliably bring about the guiding aim, in contrast with those of the novice. In basic episodes of joint attention, it plausible that this reliability claim holds true: basic episodes of joint attention consist of emotive communication about the relevant object as well as mutually responsive experiences of attention to that object. In the context of such an interaction, it is plausible that a participant’s activity reliably brings about the content of the social need: alongside gaze-following, the sharing of emotive states about the relevant object allow the participants to reliably make sense of what one another is doing. In addition, as mentioned in the previous subsection, the sharing emotive states, driven by the social motivation, is a type of communication that need not involve concep-

²¹Here I follow the structure of Dickie’s (2015) discussion of coordination in lone thought: she is concerned with *intrapersonal* coordination, in contrast to the present focus on *interpersonal* coordination. Dickie claims that the immediate justification one often has for taking one’s beliefs to be of the same object is practical, given that the activity of forming such beliefs is driven by a need to think about ordinary objects. But note that the correctness of her account of the aboutness of individual thought is orthogonal to the present topic. But, in following the structure of her account, the two ways in which the discussion to come is similar to hers are as follows. First, I also explain epistemic phenomena by likening them to more paradigmatic cases of rational action. Second, in doing so I consider two competing refinements of the general notion of practical justification, one broadly naturalist and externalist and the other broadly non-naturalist and internalist, and ultimately endorse the latter.

tual resources. And it is also plausible that the capacities of gaze-following neither need involve conceptual resources. Hence, the reliability component does not threaten the possibility of direct coordination.

A simple externalist account of practical justification claims that this reliability component exhausts practical justification. According to this view, an activity has RELIABILIST PRACTICAL JUSTIFICATION in virtue of it being a reliable way of satisfying the content of its guiding motivational state. The proposal that reliability exhausts practical justification is simple and hence appealing. Yet I have already given reason for doubting it. Compare again eating food in the regular way with, for instance, somehow directly placing food into one's stomach. Eating food in the regular way has a special status that other ways of reliably fulfilling the content of one's hunger do not; and that difference in special status is explained if only some ways of reliably satisfying the content of a guiding motivational state have practical justification.

Thus, there is a component to practical justification in addition to the reliabilist one. Now, the insight behind an internalist view of justification is that reliability is not sufficient. With regard to theoretical justification, the subject must have a reason for the belief. A reason may in some cases be a little argument one holds inside their head. But that cannot be all cases, on pain of overintellectualization. In simple cases, the internalist component may be satisfied by the phenomenological character of the relevant mental states (Campbell 2002, Smithies 2011, 2019, Dickie 2015). From the practical perspective, what the subject is doing must make sense to her, for which phenomenology may play a role. Consider a simple prehension: you are thirsty so reach out for your bottle of water on the table in front of you. Your experience of the bottle plays a role in *explaining*—in more than a mere causal sense—the way your arm moves in grasping it; your experience of the bottle as being where it is does not simply cause your arm to move in the right direction, but also illuminates to you as the agent why it should do so. In addition, your thirst and its consequent reduction play roles in explaining that same activity: your thirst is part of *why* you act in that way—in more than a mere causal sense—and the action only makes sense in that light if it reduces that feeling. Or, in our running example, only the ways of reliably satisfying the content of one's hunger that are appropriately related to its motivational phenomenology of pressure and release have practical justification. Generalizing, an activity has PHENOMENOLOGICAL PRACTICAL JUSTIFICATION in virtue of it (i) being a reliable way of satisfying the content of its guiding motivational state, but also (ii) bearing an appropriate relation to the phenomenology of the guiding motivational state as well as the relevant experiences.²²

Accordingly, both perceptual and motivational phenomenology play a role in accounting for the practical justification had by the activity of a participant in a basic episode of joint attention. The total phenomenological character of such an episode involves perceptual phenomenology from the experience of attending to the object and the other, and motivational phenomenology

²²This account might need to be generalized further, so that the internalist component may also be satisfied by beliefs, since beliefs are also sometimes rationally related to action. For instance, you believe that collecting possessions is the key to happiness, so every day you go to the mall. But I set aside beliefs since they are not relevant to the simple activity relevant to the present topic.

from the social motivational state. Recall now Campbell's perceptualism, on which the total phenomenological character of an episode of joint attention is exhausted by perceptual phenomenology. That account of joint attention is a simple extension of Campbell's approach to lone attention: in joint attention, perceptual phenomenology is like a window that makes other minds transparent alongside an external world of objects and their properties. So on his view there is a parallel between the basic way that the external world and other minds are revealed to us. But one might be unhappy with the metaphor this strong analogy generates: in our relation to social reality it is not as if we are alone, each peering at one another through our solitary windows, not noticing the empty street below.²³ And, as Eilan argues, that Campbell's view provides such a straightforward analogy leads to its failure in capturing genuinely joint attention, since there are cases of mutual covert attention where the relevant perceptual phenomenology is present but not joint attention. In contrast, it is plausible that the present proposal can meet Eilan's challenge, since both perceptual and motivational phenomenology play a role in the practical justification that grounds direct coordination, and the relevant motivational phenomenology is not present in cases of mutual covert attention.

Having shown that the direct coordination present in basic episodes of joint attention is non-conceptual, we are now in a position to address concept acquisition problem as it arises for my account. Recall that resolving that problem involves explaining how the concept of the essentially joint activity of making sense of one another may be gained. In order to provide that explanation, we must identify a way in which concepts may be gained partly in virtue of engagement in joint activity, and not merely on the basis of perceptual experience.

Recall the two important features of experiences, already emphasized, which together show how experience does provide the basis for gaining concepts. The first feature is nonconceptuality: having an experience with a certain content does not require one possess the concepts necessary for the specification of that content. This feature secures the possibility that experience may explain how one gains a concept, since having an experience does not itself require that one possess the concepts relevant for specifying its content. The second feature is rich phenomenology: what it is like to have an experience consists of a detailed view of external reality. This feature fills out a positive account of the way in which experience explains how we gain concepts on that basis: what it is like to have an experience is a rich presentation of how things are "out there".

The social need is relevantly similar to experiences in having those two features. Recall that the need is a nonconceptual motivational state. It also has a rich motivational phenomenology of pressure and release related to certain joint interactions with others. But there is one crucial qualification: this phenomenology is not transparent in the sense that perceptual phenomenology is; it does not consist of a view of external reality. Instead, that motivational phenomenology is based upon our own need to understand others and make ourselves understood. Thus, the way in which we gain the concept of the essentially joint activity of making sense of one another is by conceptualizing ourselves—our own "form of life" (Wittgenstein [1953] 2009: §19)—as we are

²³Thanks to Johannes Roessler for discussion on this point.

pushed early on by our social need into basic episodes of joint attention.

4 Conclusion

Let me conclude by first summarizing the main steps of the discussion in this paper. I began in §1 by establishing the direct coordination constraint on basic episodes of joint attention, which refined the central motivation for a perceptualist account. In §2, I systematized the communicative conception of joint attention, as recently motivated in contrast with perceptualism, and identified a version of Gricean communication that could play a sufficient role in filling out the communicative conception. In the final section §3, I began by responding to the necessity worry that Gricean communication is not present in basic episodes of joint attention, and in doing so uncovered a more basic type of communication, driven by a deep-seated social need. I then showed how appeal to the social need helped provide a non-perceptualist account of direct coordination, and also meet a related challenge to the version of Gricean communication identified in §2.

As already stressed, much of the paper has consisted of drawing together and developing related and recent ideas from philosophers and psychologists: how joint attention is a basic form of social interaction, the communicative suggestion about joint attention, and the existence of a social motivation. So let me now consolidate the view of joint attention and communication that has emerged and how it makes progress on understanding joint attention and its special role in our social lives. The overall proposal is that there are two types of communication that may sustain joint attention: the sophisticated Gricean one, and the simpler one based on the social need. Sophisticated episodes are sustained by joint-activity-based Gricean communication, in which the concept of making sense of one another is a part of the conceptual content of the relevant intentions. Basic episodes of joint attention are sustained by communication based upon the social need, in which jointness is brought about by our intentional activity, yet not in virtue of conceptual content.

The relationship between these two types of communication is as follows. Right out of the gate, we are pushed by our social need into communicative exchanges, including joint attention. We may then gain a conceptualization of ourselves and thereby gain a concept of the essentially joint activity that is deployed in the joint Gricean intentions. The nonconceptual way of establishing jointness in basic episodes makes achieving mutual intelligibility our own form of life, which we are able to discover in virtue of how it feels to act together. Contrast this picture again with the Shared Intentionality Hypothesis from the previous section. What that influential hypothesis claims is that we have a basic drive to engage in joint activity, of which communication is a slightly sophisticated form brought about in virtue of engagement in more basic forms, such as joint attention. In contrast, however, I have taken the communicative suggestion about joint attention seriously, which leads to the view that communication is a more fundamental form of joint activity than joint attention.

So, to go out on a limb, if all joint action—including, for instance, walking together—holds

partly in virtue of communication,²⁴ then my picture of joint attention may also shed light on joint action in general. The general picture is thus that there are two types of joint interaction: one driven by the social need, and another that holds partly in virtue of sophisticated intentions. Hence, an interesting feature of the resulting account is that it provides a *non-unified* account of jointness. In some cases we are drawn together in a committal way by our social need. Yet we may do things together with people who we would like nothing to do with. Consider the vicious back-and-forth of a presidential debate. These other cases hold in virtue of sophisticated intentions. Nonetheless, the committal form remains developmentally basic.

References

- Adamson, Lauren B. & Connie L. Russell. 1999. Emotion regulation and the emergence of joint attention. In Philippe Rochat (ed.), *Early social cognition*, Psychology Press.
- Battich, Lucas & Bart Geurts. forthcoming. Joint attention and perceptual experience. *Synthese*.
- Berridge, Kent C. 2004. Motivation concepts in behavioral neuroscience. *Physiology & Behaviour* 81. 179–209.
- Bratman, Michael. 1993. Shared intention. *Ethics* 104(1). 97–113.
- Bratman, Michael. 2014. *Shared agency: A planning theory of acting together*. Oxford University Press.
- Bruner, Jerome. 1975a. From communication to language—a psychological perspective. *Cognition* 3(3). 255–287.
- Bruner, Jerome. 1975b. The ontogenesis of speech acts. *J. Child Lang* 2. 1–19.
- Bruner, Jerome. 1977. Early social interaction and language acquisition. In H. R. Schaffer (ed.), *Studies in mother-infant interaction*, Academic Press.
- Bruner, Jerome. 1983. *Child's talk: Learning to use language*. Oxford University Press.
- Campbell, John. 1987. Is sense transparent? *Proceedings of the Aristotelian Society* 88. 273–292.
- Campbell, John. 2002. *Reference and consciousness*. Oxford University Press.
- Campbell, John. 2005. Joint attention and common knowledge. In Naomi Eilan, Christoph Hoerl, Teresa McCormack & Johannes Roessler (eds.), *Joint attention: Communication and other minds: Issues in philosophy and psychology*, Oxford University Press.
- Campbell, John. 2011. An object-dependent perspective on joint attention. In Axel Seemann (ed.), *Joint attention: New developments in psychology, philosophy of mind, and neuroscience*, The MIT Press.
- Campbell, John. 2018. Joint attention. In Kirk Ludwig (ed.), *The routledge handbook on collective intentionality*, Routledge.
- Carpenter, Malinda & Kristin Liebal. 2011. Joint attention, communication, and knowing together.

²⁴This proposal is suggested by Milward & Carpenter (2018) and Roessler (2020).

- In Axel Seemann (ed.), *Joint attention: New developments in psychology, philosophy of mind, and neuroscience*, The MIT Press.
- Carpenter, Malinda, Katherine Nagell & Michael Tomasello. 1998. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development* 63(4).
- Chevallier, Coralie, Gregor Kohls, Vanessa Troiani, Edward S. Brodtkin & Robert T. Shultz. 2012. The social motivation theory of autism. *Trends in Cognitive Sciences* 16(4).
- Chomsky, Noam. 1995. Language and nature. *Mind* 104. 1–61.
- Clark, Herbert H. 1996. *Using language*. Cambridge University Press.
- Dennett, Daniel Clement. 1969. *Content and consciousness*. Routledge.
- Dickie, Imogen. 2015. *Fixing reference*. Oxford University Press.
- Dickie, Imogen. 2020. Understanding singular terms. *Aristotelian Society Supplementary Volume* 94(1). 19–55.
- Dickie, Imogen & Gurpreet Rattan. 2010. Sense, communication, and rational engagement. *Dialectica* 64(2). 131–151.
- Doherty, Martin & Josef Perner. 1998. Metalinguistic awareness and theory of mind: Just two words for the same thing? *Cognitive Development* 13. 279–305.
- Dretske, Fred I. 1981. *Knowledge & the flow of information*. MIT Press.
- Egan, Frances. 2014. How to think about mental content. *Philosophical Studies* 170(1). 115–135.
- Egan, Frances. forthcoming. A deflationary account of mental representation. In Joulia Smortchkova, Krzysztof Dolega & Tobias Schlicht (eds.), *Mental representations*, Oxford University Press.
- Eilan, Naomi. 2005. Joint attention, communication, and mind. In Naomi Eilan, Christoph Hoerl, Teresa McCormack & Johannes Roessler (eds.), *Joint attention: Communication and other minds: Issues in philosophy and psychology*, Clarendon Press.
- Eilan, Naomi. 2015. Joint attention and the second person. Draft.
- Eilan, Naomi. 2018. Knowing and understanding other minds: On the role of communication. Manuscript.
- Eilan, Naomi. 2020. Other i's, communication, and the second person. *Inquiry*.
- Evans, Gareth. 1982. *The varieties of reference*. Clarendon Press.
- Evans, Gareth. 1985. Understanding demonstratives. In *Collected papers*, Clarendon Press.
- Fagin, Ronald, Joseph Y. Halpern, Yoram Moses & Moshe Vardi. 1995. *Reasoning about knowledge*. The MIT Press.
- Fine, Kit. 2007. *Semantic relationism*. Blackwell Publishing Ltd.
- Fine, Kit. 2012. The structure of joint intention. Unpublished Manuscript, available on academia.edu.
- Flavell, John H. 1977. The development of knowledge about visual perception. In C. B. Keasey (ed.), *The nebraska symposium on motivation: Social cognitive development*, vol. 25, 43–76.
- Flavell, John H. 1992. Perspectives on perspective taking. In H. Beilin & P. B. Pufall (eds.), *Piaget's*

- theory: Prospects and possibilities*, vol. 14 The Jean Piaget Symposium Series, 107–139.
- Flavell, John H., Barbara A. Everett, Karene Croft & Eleanor R. Flavell. 1981. Young children's knowledge about visual perception: Further evidence for the level 1–level 2 distinction. *Developmental Psychology* 17(1). 99–103.
- Frankfurt, Harry G. 1978. The problem of action. *American Philosophical Quarterly* 15(2). 157–162.
- Frege, Gottlob. 1960. On sense and reference. In Peter Geach & Max Black (eds.), *Translations from the philosophical writings of gottlob frege*, Basil Blackwell.
- Gilbert, Margaret. 1990. Walking together: A paradigmatic social phenomenon. *Midwest Studies in Philosophy* 15.
- Gilbert, Margaret. 2013. *Joint commitment: How we make the social world*. Oxford University Press.
- Grice, Paul. [1957] 1989. Meaning. In *Studies in the way of words*, Harvard University Press.
- Grice, Paul. [1969] 1989. Utterer's meaning and intentions. In *Studies in the way of words*, Harvard University Press.
- Harman, Gilbert. 1990. The intrinsic quality of experience. *Philosophical Perspectives* 4. 31–52.
- Harris, Daniel. 2019. We talk to people not contexts. *Philosophical Studies* .
- Heck, Richard G., Jr. 2000. Nonconceptual content and the "space of reasons". *The Philosophical Review* 109(4). 483–523.
- Heck, Richard G., Jr. 2007. Are there different kinds of content. In Brian P. McLaughlin & Jonathan D. Cohen (eds.), *Contemporary debates in philosophy of mind*, Blackwell.
- Hobson, R. Peter. 2002. *The cradle of thought*. Pan Macmillan.
- Hobson, R. Peter. 2005. What puts the jointness in joint attention? In Naomi Eilan, Christoph Hoerl, Teresa McCormack & Johannes Roessler (eds.), *Joint attention: Communication and other minds: Issues in philosophy and psychology*, Oxford University Press.
- Jankovic, Marija. 2014. Communication and shared knowledge. *Philosophical Studies* 169. 489–508.
- Lederman, Harvey. 2017. Two paradoxes of common knowledge: Coordinated attack and electronic mail. *Noûs* 1–25.
- Lederman, Harvey. 2018. Uncommon knowledge. *Mind* 127(508). 1069–1105.
- León, Felipe. 2021. Joint attention without recursive mindreading: On the role of second-person engagement. *Philosophical Psychology* 34(4). 550–580.
- Lewis, David. 1969. *Convention: A philosophical study*. Harvard University Press.
- Marr, David. 1982. *Vision: A computational investigation into the human representation and processing of visual information*. W. H. Freeman and Company.
- McDowell, John. 1994. *Mind and world*. Harvard University Press.
- Mesman, Judi, Marinus H. van Ijzendoorn & Marian J. Bakermans-Kranenburg. 2009. The many faces of the still-face paradigm: A review and meta-analysis. *Developmental Review* 29(2). 120–162.
- Milward, Sophie J. & Malinda Carpenter. 2018. Joint action and joint attention: Drawing parallels between the literatures. *Soc Personal Psychol Compass* 12. 1–11.
- Moll, Henrike & Andrew N. Meltzoff. 2011. Joint attention as the fundamental basis of under-

- standing perspectives. In Axel Seemann (ed.), *Joint attention: New developments in psychology, philosophy of mind, and neuroscience*, The MIT Press.
- Moll, Henrike, Andrew N. Meltzoff, Katharina Merzsch & Michael Tomasello. 2013. Taking versus confronting visual perspectives in preschool children. *Developmental Psychology* 49(4). 646–654.
- Neale, Stephen. 1992. Paul grice and the philosophy of language. *Linguistics and Philosophy* 15. 509–559.
- Nyström, Pär, Emilia Thorup, Sven Bölte & Terje Falck-Ytter. 2019. Joint attention in infancy and the emergence of autism. *Biological Psychiatry* 86. 631–638.
- Peacocke, Christopher. 1992. *A study of concepts*. MIT Press.
- Peacocke, Christopher. 2005. Joint attention: Its nature, reflexivity, and relation to common knowledge. In Naomi Eilan, Christoph Hoerl, Teresa McCormack & Johannes Roessler (eds.), *Joint attention: Communication and other minds: Issues in philosophy and psychology*, Oxford University Press.
- Rakoczy, Hannes, Delia Bergfeld, Ina Schwarz & Ella Fizke. 2015. Explicit theory of mind is even more unified than previously assumed: Belief ascription and understanding aspectuality emerge together in development. *Child Development* 86(2). 486–502.
- Rochat, Philippe & Tricia Striano. 1999. Social-cognitive development in the first year. In Philippe Rochat (ed.), *Early social cognition*, Psychology Press.
- Roessler, Johannes. 2005. Joint attention and the problem of other minds. In Naomi Eilan, Christoph Hoerl, Teresa McCormack & Johannes Roessler (eds.), *Joint attention: Communication and other minds: Issues in philosophy and psychology*, Clarendon Press.
- Roessler, Johannes. 2020. Plural practical knowledge. *Inquiry* .
- Scaife, Michael & Jerome Bruner. 1975. The capacity for joint visual attention in the infant. *Nature* 253. 265–266.
- Schiffer, Stephen. 1972. *Meaning*. Oxford University Press.
- Seemann, Axel. 2010. The other person in joint attention: A relational approach. *Journal of Consciousness Studies* 17(5–6). 161–182.
- Seemann, Axel. 2019. *The shared world: Perceptual common knowledge, demonstrative communication, and social space*. The MIT Press.
- Siposova, Barbora & Malinda Carpenter. 2019. A new look at joint attention and common knowledge. *Cognition* 189. 260–274.
- Smithies, Declan. 2011. What is the role of consciousness in demonstrative thought? *Journal of Philosophy* 108(1). 5–34.
- Smithies, Declan. 2019. *The epistemic role of consciousness*. Oxford University Press.
- Sperber, Dan & Deirdre Wilson. 1986. *Relevance: Communication and cognition*. Blackwell Publishers.
- Stern, Daniel N. 1985. *The interpersonal world of the infant: A view from psychoanalysis and developmental psychology*. Harvard University Press.
- Strawson, Peter F. 1964. Intention and convention in speech acts. *Philosophical Review* 73(4). 439–

- Strawson, Peter F. 1988. Perception and its objects. In Johnathan Dancy (ed.), *Perceptual knowledge*, Oxford University Press.
- Striano, Tricia & Evelin Bertin. 2005. Coordinated affect with mothers and strangers: A longitudinal analysis of joint engagement between 5 and 9 months of age. *Cognition and Emotion* 19(5). 781–790.
- Tenenbaum, Sergio. 2015. Representing collective agency. *Philosophical Studies* 172. 3379–3386.
- Toates, Frederick. 1986. *Motivational systems*. Cambridge University Press.
- Tomasello, Michael. 1999. *The cultural origins of human cognition*. Harvard University Press.
- Tomasello, Michael. 2009. *Why we cooperate*. The MIT Press.
- Tomasello, Michael. 2014. The ultra-social animal. *European Journal of Social Psychology* 44. 187–194.
- Tomasello, Michael. 2018. How children come to understand false beliefs: A shared intentionality account. In *Proceedings of the national academy of sciences of the united states of america*, .
- Tomasello, Michael. 2019. *Becoming human: A theory of ontogeny*. Harvard University Press.
- Tomasello, Michael, Malinda Carpenter, Josep Call, Tanya Behne & Henrike Moll. 2005. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28. 675–735.
- Trevarthen, Colwyn. 1979. Communication and cooperation in early infancy: a description of primary intersubjectivity. In M. Bullowa (ed.), *Before speech: The beginning of human communication*, 321–347. Cambridge University Press.
- Tronick, Edward, Heidelise Als, Lauren Adamson, Susan Wise & T. Berry Brazelton. 1978. The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child Psychology* 17(1). 1–13.
- Wittgenstein, Ludwig. [1953] 2009. *Philosophical investigations*. Wiley-Blackwell.