

Opening up the openness of joint attention

Lucas Battich

Faculty of Philosophy, Ludwig-Maximilians-Universität München

Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität München

Email: lucas.battich@campus.lmu.de

June 2021

Abstract

The ability to engage in joint attention, in which two individuals attend to the same object or event together, is considered fundamental for language learning, for understanding others and for joint actions. Joint attention is often defined as a mutually open, or transparent relation between co-attenders. But how should this openness be characterised? Two broad theoretical views have been proposed. One view reductively accounts for the mutual awareness characteristic of joint attention in terms of individual mental states and properties. In contrast, according to non-reductive views, mutual awareness is based on some primitive intersubjective relation, which is irreducible to the individual states of each co-attender. I argue that tensions in these approaches arise from the attempt to address both normative and cognitive explananda simultaneously. Both approaches are primarily designed to tackle the normative epistemological concerns of joint attention, and their problems arise when they conflate these concerns with psychological ones. Drawing from evidence in developmental and cognitive psychology, I outline the case for a cognitive-first account of joint attention based on a weaker notion of openness and mutual awareness. I conclude by assessing the epistemic implications of this account.

Keywords: Joint Attention, Common Knowledge, Mutual Awareness, Social Cognition

“I think (and I think Hume did too) that, insofar as it’s about the analysis of justification and the like, epistemology hasn’t really got much to do with psychology.” (Fodor, 2003: 4)

1. Introduction

Two people sit at a table with a piece of cake between them. They look at the cake, exchange glances and smile. They are thus both attending to the cake. Importantly, they are both at the same time aware of each other’s attention. If one of them was to say “Grandma made it”, the referent would be clear to both of them. This is a paradigmatic case of joint attention.

It is widely agreed that joint attention, in which two individuals attend to the same object or event together, plays an important role in language development and communication, in joint action, and in the progressive understanding that others can have different perspectives than our own. More generally, joint attention supports the development of mentalising, the ability to comprehend other people’s mental lives (Carpenter et al., 1998; Moore & Dunham, 1995; Mundy, 2018). Within philosophy, joint attention has been considered essential for the ability to distinguish self from other (Bermúdez, 1998), the constitution of a common ground for communication (Clark, 1992), and the concept of a shared objective world, where objects are attended in common (Campbell, 2011; Davidson, 1999). However, there is scant agreement on how joint attention should be analysed and how it is involved in all these capacities.

This paper aims to contribute to theoretical research on joint attention by examining the notion of a shared “openness” between co-attenders. Most researchers agree that a key feature of the state of joint attention is that it is public or “out in the open” among the co-attenders. It is fully and immediately “transparent”, or “mutually manifest”, to them that they are jointly attending to the same object or state of affairs (Bakeman & Adamson, 1984; Calabi, 2008; Campbell, 2011; Carpenter & Liebal, 2011; Eilan, 2005; Peacocke, 2005; Tomasello, 1995). This feature puts the jointness in joint attention and distinguishes it from parallel attention, where we attend to the same object unilaterally, completely unaware of each other. In joint attention, co-attenders are *mutually aware* of their shared attention to the same target. But how should this notion of mutual awareness be analysed? Two broad theoretical views have been proposed. One view reductively accounts for the mutual awareness characteristic of joint attention in terms of individual mental states and properties. In contrast, non-reductive views hold that mutual awareness is based on some primitive intersubjective relation, which is irreducible to the individual states of each co-attender.

I have two goals in this paper. First, I argue that much of the debate between reductive and non-reductive views arise through the conflation of two distinct explanatory aims: the aim of explaining the normative role of joint attention in justifying joint endeavours, and the cognitive aim of explicating the psychological capacities involved in joint attention.

Reductive and non-reductive approaches are primarily concerned with the normative aim, and their problems arise when they extend their scope to tackle the cognitive aim (section 2). Both approaches, I argue, turn out to be conceptually equivalent when the focus is strictly normative. Second, I outline the case for a cognitive-first account of joint attention based on a weaker notion of mutual awareness, in line with empirical evidence from developmental psychology (section 3). I suggest that mutual awareness in joint attention is not something co-attenders must arrive at, but that it is often implicitly assumed, and that we must (un)learn that other people may not attend to the same things we attend. I conclude by returning to the normative question, assessing how this proposal can address the epistemic justificatory role of joint attention in our social lives. I suggest differentiating between a *psychological* notion of joint attention and a *normative* notion. The proposed account, I argue, shifts the normative question from explaining how individuals can attain a justified and well-founded relation of mutual awareness, to explaining how individuals abstain from defaulting to possibly erroneous assumptions of mutual awareness.

2. Assessing the reductive and non-reductive views

2.1. Explanatory aims

It is commonly assumed that a primary functional role for joint attention is to provide a rational basis for further coordinated activities between agents. Joint attention is thus considered a necessary but not sufficient condition for individuals to engage in collective activities (Brinck & Gärdenfors, 2003); and plays an essential role in the constitution of the common ground involved in communication (Bruner, 1995; Clark, 1992; Tomasello, 2008). John Campbell shares in the current consensus by associating the openness of joint attention with this functional role:

[W]hatever else is true of it, joint attention has an “openness” about it — there’s some sense in which the situation is “open” to both attendees in a case of joint attention — in virtue of which joint attention ordinarily plays a distinctive role in rational, coordinated action. (Campbell, 2011: 417)

On the other hand, joint attention is a psychological phenomenon, and any theoretical analysis will be constrained by psychological plausibility. For example, Christopher Mole (2017) notes that, “since joint attention is achieved by young children, its achievement cannot plausibly be thought to make any sophisticated intellectual demands” (cf. Roessler, 2005). We need to spell out what the mental and cognitive requirements for joint attention are: how is openness achieved? An account of the openness of joint attention is therefore seen as a necessary step in addressing these two different explananda:

- i. Normative: How does joint attention rationally support joint beliefs and joint actions, and yields shared knowledge of our environment?

- ii. Cognitive: What cognitive capacities and mental processes or understanding are involved in joint attention, which even infants and young children may be capable of having?

Though they needn't be, these two questions are often tackled at once in the current debate. Supposedly, the openness of joint attention must be conceptualised in such a way that it provides a normative epistemic basis for rational behaviour regarding the co-attended object, and simultaneously shed light on the mental wherewithal necessary to achieve said openness. I examine the contrast between reductive and non-reductivist accounts of joint attention in light of these two different explanatory aims. Both accounts take the normative question (i) as a starting point and motivation. It is usually assumed that the functional role of joint attention is to provide the categorical grounds that allow subjects to engage rationally in joint projects concerning the jointly attended object, including communicative speech referring to the object. In the rest of this section, I show how the problems for both accounts arise only when they are taken to simultaneously address the cognitive question (ii).

2.2. The reductive approach

According to the reductive approach, the openness in joint attention can be explained in terms of the individual mental states of each co-attender. While this view does not deny the experiential or phenomenal character of joint attention, it holds instead that this character is better accounted for in terms of individual epistemic states. Reductive views tend to analyse the openness of joint attention in terms of common knowledge or similar notions such as mutual awareness, constructed in line with the analyses from Lewis (1969) and Schiffer (1972). On this definition, A and B are jointly attending to x when they are mutually aware, or enjoy common knowledge, that each of them is attending to x . Common knowledge is defined as giving rise to iterative structures like the following:

p is common knowledge between A and B iff

- A knows that p ,
- B knows that p ,
- A knows that B knows that p ,
- B knows that A knows that p ,
- and so on, *ad infinitum*.

On this construal, joint attention involves nested psychological states that both A and B would have to entertain about each other. This is the hallmark of recursive mindreading, where one subject (A) attributes mental states to another (B knows that ...), which, in turn, refer to the first subject's own mental states (B knows that A knows that ...). Most criticisms of this approach centre on the requirement of recursive mindreading (León, 2021). One argument from phenomenology notes that the openness in joint attention is immediate

and effortless, so that it is implausible that joint attention requires recursive mindreading (e.g. Gallagher, 2011). Related arguments point to the computational complexity of recursive mindreading, arguing that it is cognitively demanding and psychologically implausible even for adults (Campbell, 2018; Eilan, 2005).

Felipe León (2021) has accurately summarised the problems of recursive mindreading. My point, however, is that the claim that reductive accounts necessitate recursive mindreading already conflates normative and cognitive explananda. To tease out how these differ, I focus on a key argument against the reductive approach, which appeals to “coordinated-attack” scenarios (Campbell, 2005, 2018; Wilby, 2010).¹

Coordinated-attack scenario: Two individuals in separate booths must both attack the same target among many, at the same time. For this, they will have to coordinate their individual actions. Additionally, there is always a non-zero chance of distorted communication between booths, so that when one individual has chosen a target and communicates this to the other, they will not know for sure whether the communication has been received. Supposing that the second individual does receive the message specifying the target, they could, in turn, send a message back to confirm receipt of the target. But, again, they will not be sure whether this confirmation has been received. To make a coordinated attack rational, both individuals will have to go through an infinite iteration of messages confirming the preceding confirmation.

Coordinated-attack scenarios are often assumed to show that no finite iteration of inferences will allow the participants to engage in rational coordinated attack and secure victory. Yet in normal situations, where we are both present in the same physical space, we can easily arrive at a successful coordinated outcome. When both individuals and the target are all co-present, as in most normal circumstances, then “everything is out in the open to such an extent that we can rationally attack” (Campbell, 2005: 292).

For Wilby (2010), the coordinated-attack problem “highlights that once one gets embroiled in *supposing* that an act of transparent communication or shared knowledge requires a set of hierarchical to-ing and fro-ing about who knows what, then there will be no end to the matter” (my emphasis). Going up only two or three levels up the recursive chain will not suffice. As Wilby (2010) and Campbell (2018) note, for any level in the chain, that level does not lead to openness or mutual awareness, or else will require a further step in

¹ The coordinated-attack problem was first introduced in its current form by Akkoyunlu et al. (1975) in the context of dynamical systems engineering and, while it soon became a key fixture in epistemic logic, it had little to do with mental representations and psychological processes.

the recursive chain². The conclusion of the coordinated-attack argument is what Wilby calls a disastrous “paradox” for the reductive approach: mutual awareness requires an infinite recursion of overlapping mental states, but this requirement is psychologically implausible.

There are two points where the argument from coordinated-attack fails. First, there is nothing in the coordinated-attack scenario that implies that the iterations must be actually *represented* in the mind of each individual in a case of joint attention. Presupposing that each iteration must be represented already relies on the convergence of both normative and cognitive aims described above. Second, this rendering of the “paradox” assumes that, on the iterative approach, common knowledge is the *result* or *endpoint* of the iterations. On this assumption, the openness in normal joint attention scenarios would be equivalent to the openness which the individuals in our separate booths are infinitely pursuing. This would spell problems for the reductionist: if a chain of iterations does not produce common knowledge in the coordinated-attack scenario, it cannot explain how the openness of joint attention is achieved. Yet the reductionist is not committed to making such assumption in the first place. The view that a reductive approach to joint attention involves recursive mindreading is based on normative analyses of common knowledge such as those by Lewis and Schiffer. Traditionally, their approach presumes that there are situations with some *finite* conditions, out of which the infinite iterations logically follow. What makes a situation a common knowledge situation are those finite conditions, not the recursive iterations themselves, nor their result or endpoint. In other words, the iterations arise as logical implications which follow from some given finite situation, and which are not necessarily represented in anyone’s reasoning (Lewis, 1969: 53). Coordinated-attack scenarios only show that the individuals in their separate booths lack the appropriate finite conditions.

To date, providing a good account of those finite conditions has been proved somewhat problematic, but only when, in addition, such conditions must also account for the psychological processes and mental wherewithal necessary for joint attention or common knowledge. If we do away with this cognitive explananda, we are left only with the normative question. Under the normative aim, the openness of joint attention can be seen as a purely normative epistemological notion, and there is nothing psychological about it. Analyses of joint attention and mutual awareness in the style of Schiffer’s and Lewis’ are not necessarily committed to recursive mindreading or any regression of mental states.

² One may suggest that people just reason two or three levels up the hierarchy and then conclude their knowledge is shared. This is an empirical, not a logical, suggestion; and one with no evidential support (e.g., Liddle & Nettle, 2006; Thomas et al., 2014). The approach implicit in this suggestion is normative-first: start with a logical, normative theory, and truncate it to fit what (we think) humans can do. I suggest adopting instead a cognitive-first approach and evaluate any normative implications thereafter.

They are not in principle committed to any view about psychological processes at all, and so assuming that they must involve the performance of an infinite chain of mental states is a misapplication of a normative analysis into a psychological straitjacket.³

This conclusion leaves unanswered what mutual awareness is, psychologically speaking, as a mental state enjoyed by both adults and infants. In other words, the upshot is that we are giving up the cognitive explanandum of joint attention.

2.3. The non-reductive approach

Prominent non-reductive approaches tend to focus on the experiential aspects of joint attention (Campbell, 2005, 2018; Seemann, 2019; Wilby, 2010). It is in virtue of its phenomenal character that joint attention plays an epistemic role in justifying shared beliefs and joint activities. On this approach, there is a primitive intersubjective relation behind joint attention, in the sense that it cannot be analysed any further in terms of its individual components. John Campbell thus proposes that joint attention is a primitive type of conscious state: “just as the object you see can be a constituent of your experience, so too it can be a constituent of your experience that the other person is, with you, jointly attending to the object” (2005: 288, cf. 2018). Naomi Eilan argues that joint attention is grounded in experiences of “you-awareness” and “communication-as-connection”, which are primitive conscious states (2015). Following Campbell, Axel Seemann (2019) argues that our perceptual experience during joint attention is a primitive joint state. What each of us experiences during joint attention cannot be reduced to our individual psychological states, but is determined by the spatial organisation between us and the common object of our attention (Seemann, 2019: 75).

Campbell’s analysis, in particular, is based on the premise that joint attention can be explained fully in terms of perceptual experience, and thus is not susceptible to an explanation in terms of the knowledge, beliefs, or awareness of the two participants (Campbell, 2018: 120). For this reason, it has proved to be an attractive theoretical position for cognitive and developmental psychologists (e.g. Moll & Meltzoff, 2011; Hobson & Hobson, 2011). One serious criticism of this approach is that it simply embeds in the analysis of openness or mutual awareness the property that is to be explained, i.e. the openness of joint attention (Peacocke, 2005: 300). Further problems arise due to the narrow focus on perceptual experience. Since the non-reductive approach allows for dissociations between one’s perceptual experience and one’s beliefs about that experience, it also allows for the possibility that one is engaged in joint attention but mistakenly believes that is not — a conclusion that goes against the key feature of joint attention that the view attempts

³ For a purely normative treatment of phenomena that share the openness under consideration, particularly the common ground between interactants and its role in communication, see Geurts (2019).

to explain: co-attenders are mutually aware of their joint attention (Battich & Geurts, 2020).

As it is perhaps already evident, the key motivation for non-reductive approaches is to address the normative explananda of joint attention. Its limitations arise when it is also taken to provide insight into the cognitive question. When focusing exclusively on the normative question, however, the openness of joint attention becomes merely a description of phenomenal aspects of experience. Joint attention rationalises joint action and joint beliefs in virtue of the two co-attenders being in the appropriate phenomenal state. Being primitive, this state cannot be further explained. But neither does it inform us about the psychological capacities behind joint attention. On the non-reductive approach, joint attention is treated from a third-person point of view, as it merely asks whether an external *ascription* that A and B jointly attend to x is true. A description of a primitive intersubjective relation between co-attenders can make such ascriptions true, but this description will be entirely silent on the psychological states of each individual (cf. Schmitz, 2015: 239). The upshot, of course, is that we are giving up the cognitive explanandum (aim ii).

2.4. Conceptual equivalence between iterative and primitivist approaches

For primitivists like Campbell (2005) and Seemann (2019), the state of mutual awareness in joint attention is a factive state. You cannot be aware that the other person is currently co-attending with you to the same perceptual target, unless that person is, in fact, a co-attender. Of course, you could be wrong about your co-attendance, but then you would not be in a state of joint attention. The factive character of the mental states of each co-attender is also usually assumed for the reductive, iterative approach. Two people are mutually aware of p only when both of them are equally justified to follow the infinite logical implications of their joint epistemic state. Interestingly, the assumption that the mental states of each co-attender are factive implies that the iterative and primitivist views are conceptually equivalent — at least under some versions of each view. In particular, the equivalence holds for Schiffer’s analysis of common knowledge, commonly taken as a paradigm of the reductive, iterative approach when applied to the openness in joint attention. The brilliance of Schiffer analysis is that it proposes a *finite basis* for common knowledge, out of which the iterations would follow logically. Therefore, it does not necessitate an infinite recursion of mental states, to be represented in the minds of each individual. On Schiffer’s analysis, A and B mutually know that p iff there are properties F and G such that:

1. A is F.
2. B is G.
3. Both being F and being G are sufficient for knowing that p , that A is F, and that B is G.

4. For any proposition q , if both being F and being G are sufficient for knowing that q , then both being F and being G are sufficient for knowing that both being F and being G are sufficient for knowing that q . (Schiffer, 1972: 34-5)

Given this finite base, the infinite number of iterations characteristic of common knowledge can be generated by feeding (3) to the recursive clause in (4), and reapplying (4) to each new result over and over. This analysis relies heavily on the generating properties F and G, which, Schiffer proposes, refer to the property of being “a visibly ‘normal’, open-eyed, conscious person”:

If a “normal” person (i.e. a person with normal sense faculties, intelligence, and experience) has his eyes open and his head facing an object of a certain size (etc.), then that person will see that an object of a certain sort is before him. (Schiffer, 1972: 31)

Moreover, all people know that normal people will behave in this way, and they can easily tell when someone is normal (Schiffer, 1972: 33). For this analysis to work, the property of “being normal”, however, must be relativised to the specific situation in which the co-attenders currently are (Wilby, 2010). Some situations will require, for example, that assumptions about normal hearing, rather than normal sight, be included in the normality properties F and G to allow for common knowledge towards an auditory event in the environment. Importantly, Wilby (2010) has shown that F and G are more intimately related than Schiffer initially presumed. Schiffer’s four clauses, in particular (3), together with the facticity assumption (which makes knowledge a factive state by definition: if X knows p , then p is true), imply that the two generating properties F and G are necessary and sufficient conditions for each other: A is F iff B is G (for the deductive proof, see Wilby, 2010: 91).

Given the biconditional relation between F and G, these properties can be logically replaced by a primitive relational property H, so that H iff F and G. Thus, Schiffer’s analysis of common knowledge turns out to be logically equivalent to an analysis including the single intersubjective property H, where A and B are both H, and each of them can only be H when the other person is likewise H. But now the analysis includes a primitive intersubjective element irreducible to the mental states and properties of each individual. Schiffer’s analysis of common knowledge is conceptually identical to a primitivist analysis. Common knowledge, under Schiffer’s analysis together with the facticity assumption, is a relational state irreducible to the individual cognitive states of the individuals in question (Wilby, 2010: 92). It is important to be clear about the implications of this equivalence. Wilby suggests that the equivalence, together with the psychological implausibility of the iterative approach, arbitrate in favour of the primitivist approach to common knowledge and related epistemic notions, such as mutual awareness and joint attention. This suggestion is unwarranted, however. Wilby’s argument demonstrating the equivalence between the two views constitutes a redrawing of the normative aspects of common knowledge. In principle, the primitive relational property H has little to do with psychology and the actual mental states of an individual (Wilby, 2010: 93, admits as much). One could go further, of

course, and interpret H as an irreducible joint *psychological* state. But this interpretation only brings back the limitations of the primitivist approach: the openness of joint attention becomes an irreducible psychological state, and, as Wilby himself notes, it is suspect how much explanatory work such irreducible notion can play in psychological theories and experimental research. In particular, the cognitive explananda is left untouched: what mental processes and understanding are involved in joint attention, which even infants and young children may be capable of having? This defeatist outcome is due, in part, to the equivalence relying on the facticity assumption. While this clause is commonly assumed for knowledge (Williamson, 2000), it arguably does not hold for psychologically determined states such as beliefs and awareness. It is dubious that a person's *psychological* sense of being in a situation of joint attention leads, as a matter of logical necessity, to joint attention being true. We have to allow for the possibility that someone can take themselves to be engaged in joint attention while they are, in fact, not.

The conceptual equivalence of iterative and primitivist analyses holds when both analyses are strictly considered as epistemic normative theories. Their unreconcilable differences arise only when they are taken to address, in addition, the cognitive question regarding the psychological processes behind joint attention. If we are merely interested in the normative dimension of joint attention, both reductive and non-reductive approaches to joint attention seem equally valid. However, if we are interested in the psychological states of real humans and children, and not just in the rational states of epistemic agents, then we should acknowledge the limitations of such "normative-first" analyses of joint attention. Neither of the two approaches makes it possible to address the cognitive instead of the normative question.

3. A cognitive-first approach to joint attention

3.1 Mutual awareness is assumed

In 2003 Michael Tomasello remarked that child language acquisition is not a logical problem, but an empirical one. He urged that a theory of human linguistic competence should be based less on analogies to formal languages, and more on empirical research in the cognitive sciences (2003: 328). Unfortunately, Tomasello himself didn't fully apply this dictum to the topic of common knowledge and joint attention, assuming, along with many others, that it must involve either something akin to recursive mindreading, or some primitive non-analyzable plural "we" subject, which is in turn presumably produced by as yet undiscovered unconscious mechanisms (cf. Zawidzki, 2013):

From early on as well, infants communicate with others referentially, inviting them to jointly attend to something, and this requires recursive inferences about mental states embedded in mental states. (Tomasello, 2019: 44)

Given the narrow "normative-first" focus of traditional approaches to joint attention, however, I propose that, if we are interested in the cognitive question, there are no strong

reasons to assume a priori that the key feature of joint attention is a fully symmetric epistemic or phenomenal openness between co-attenders. Instead, I propose that a more fruitful approach to describe the triadic interaction of joint attention is to concentrate on what factors or aggregate of factors each individual co-attender is responding to, so that this interaction can be established, without presuming in advance the nature of the epistemic or phenomenal sophistication they must achieve. In this section, I provide the outline of a “cognitive-first” approach to assess the jointness of joint attention. The starting point is to leave normative concerns on the side for the time being. The aim of a cognitive-first approach is not to arrive at a *justification* for the mental states of an individual during joint attention. We are not (yet) concerned with the rationality of those mental states.

In a nutshell, I suggest that mutual awareness in joint attention is neither a primitive nor reductive intersubjective relation that co-attenders must arrive at, but that it is often implicitly assumed, and that we must (un)learn that other people may not attend to the same things we attend, or may not share the same perceptual knowledge we are currently enjoying. This view is supported by research in developmental and cognitive psychology. Two-year-old children typically assume that an adult interacting with them will share their perceptual perspectives (Epley et al., 2004; Moll & Meltzoff, 2011). In one particular experiment, two-year-old children shared visual attention of two objects with an adult, one by one. The child was then presented with a third object, which the adult could not see. In one condition, this was because the adult was present behind a barrier, so that they didn’t have visual access to the third object but continued to communicate verbally while the child inspected the object. In another condition, the adult was completely absent from the room when the third object was shown to the child. The task was to identify which of the three objects was new for the adult when she explicitly requested to the child for the “one she has not seen before”. Children were able to correctly select the new object when the adult had been absent from the room, but they were not able to differentiate between new and old objects when the adult was behind the visual barrier and could not see the object but still engage with them verbally (Moll et al., 2011). Moll and colleagues interpret these results as showing that physical co-presence and some form of minimal engagement is enough for children to assume that they are sharing their perceptual experiences with the adult (see also Hobson & Hobson, 2011). These findings are in line with, for example, the everyday experience of a child talking on the telephone and assuming that the other person is aware of what they are pointing to. The impulse to assume openness or sharing experiences or knowledge is not restricted to children alone. A similar phenomenon is observed in adults, where someone’s own knowledge will affect, however implicitly, their ability to reason about another person’s beliefs (Epley et al., 2004). Referred to as the “curse of knowledge”, people are egocentrically biased to assume that others know what they themselves know (Birch & Bloom, 2007; Farrar & Ostojić, 2018).

On the view I propose, then, to jointly attend to an object is to pre-reflectively assume that one’s attention to the same object is shared with someone else’s. Contrary to the traditional reductive view, attaining perceptual common knowledge towards the same object is not cognitively taxing, but curbing it down is: taking into consideration whether other

people do not share your object of attention is cognitively demanding, at least during development and in novel situations with no prior precedents. Unlike non-reductive views which posit some intersubjective primitive phenomenon behind joint attention, the view I propose is anchored on the individual. It concerns the mental processes that an individual A must go through in order to say that she is jointly attending with B to x (and is mutually aware with B about so being in joint attention to x). Of course, B and her mental processes will often come into the picture too. However, in this analysis, B mustn't necessarily be a minded individual. A can engage in joint attention with, e.g., a computer avatar or with animals. Whatever B knows or is aware of is not constitutive of this analysis. This account concerns the mental and psychological states of A alone, so that we can say that A takes herself, from her perspective and her practical purposes (though not necessarily consciously) to be jointly attending with B to x.

This proposal allows us to get a grip on the cognitive question without as yet being misled by normative concerns. What cognitive capacities and mental processes or understanding are involved in joint attention? What minimal capacities are necessary to pre-reflectively assume that one is sharing attention to the same object with others? Based on the studies by Moll and colleagues, we can start with a set of minimal cues that I as a co-attender should be capable of recognising:

- You are a separate individual from me.
- You are physically present.
- You engage with the world as I do.
- You are interacting with me.

Once I recognise these, I can form the subpersonal representation that the attention towards x is shared. This cognitive process could be paraphrased as "I have a certain relation to x, and since you and I are so similar, and you are interacting with me, (I assume) you have it too." It becomes clear that the notion of mutual awareness I am using here is considerably weaker than the notion used by reductive and non-reductive approaches. When I assume mutual awareness, this does not imply that this assumption must be fully conscious, reflective or deliberate, or that I have to consciously entertain the proposition that you are similar to me. It does not even require having a concept of mutual awareness. The only requirement is that I recognise and respond to the cues that you provide by implicitly assuming that you engage with the object in the same way I engage to it. Such recognition and response is plausibly supported by subpersonal sensory-motor and affective processes (cf. Reddy, 2010, 2015; Gallagher, 2010). Interaction is a key component here. Without at least a minimal form of interaction (or the assumption, from my point of view, that you interact with me), it seems implausible to assume that I attend to some object *together* with someone else. Joint attention is a form of action done with others. Note that recognising these cues does not need to result in an infallible and correct assumption of openness. In Moll and colleagues' study, children interact with an adult verbally, but fail to recognise that the adult cannot visually access the same object as they do. Yet, this minimal

verbal interaction is sufficient to make the infants assume that they are jointly attending to the object, even when this assumption is, from a third-person point of view, incorrect.

Conceptual and reflective awareness need play no role to form such assumption.⁴ Of course, I could become reflectively aware of our joint engagement towards x. Usually, this occurs when the assumption of joint attention breaks or misfires. If I say to you “Grandma made it”, and you show no comprehension of the intended referent, I can become retroactively conscious that I assumed, incorrectly, that we were looking at the cake together. When interaction fails, an individual may learn to tone down their assumptions of openness in similar future interactive situations, and reevaluate the set of cues that trigger their assumption of mutual awareness in those situations.

Uncovering the set of cognitive capacities behind joint attention is at root an empirical project, not a purely conceptual one. But conceptual clarity regarding different explanatory aims can assist with this empirical project. For this reason, while the present proposal is yet underdeveloped as a full response to the cognitive question and cannot be the complete story, it serves as an illustration of a cognitive-first approach to joint attention and its openness.

3.2. Epistemic implications

Mutual awareness in joint attention is implicitly assumed. How would this approach fare in accounting for the epistemic justificatory role that joint attention is taken to play in human social lives? The normative aim is to explain how joint attention supports joint beliefs and joint actions, and yields shared knowledge of our environment. To play this role, it is assumed that the state of joint attention itself must be well-founded or justified. The reductive and non-reductive views are attempts to characterise a notion of mutual awareness in joint attention that is itself justified. The proposed account, I suggest, shifts the epistemic normative question from explaining how individuals can attain a justified and well-founded relation of mutual awareness, to explaining how individuals abstain from defaulting to possibly erroneous assumptions of mutual awareness. This approach concerns the mental and psychological states of each individual alone, so that all we can say, regarding A, is that A attends to x, and that A takes herself to be attending to x together with B. Whether A is actually justified to take herself to be in a situation of joint attention is a further question, which will depend on factors *external* to her psychology. More precisely,

⁴ In contrast to both Campbell (2005) and (2005), I do not aim to characterise the phenomenal experience of openness itself. I remain noncommittal to what the phenomenology of this assumption might be. However, it is important to note that the functional role of assuming mutual awareness, in the weaker sense used here, can be carried out without invoking its phenomenological aspect. Moreover, the present proposal is in principle compatible with interactionist and embedded embodied accounts of joint attention — as long as the analysis does not stop at some primitive intersubjective level.

it will depend on what is happening inside the mind of B: A is justified to take herself to be in a situation of joint attention with B iff B attends to x, and B takes herself to be attending to x together with A. Conversely, B is justified to take herself to be in a situation of joint attention with A iff A attends to x, and A takes herself to be attending to x together with B. We can now spell out a normative epistemic account of joint attention.

A and B jointly attend to x iff

1. A attends to x.
2. B attends to x.
3. A takes herself to be attending to x together with B.
4. B takes herself to be attending to x together with A.

This account seems at first circular. A critic may point out that conditions (3) and (4) already presuppose what we are trying to analyse, i.e. the jointness of joint attention. On the other hand, a primitivist proponent might in turn retort that (3) and (4) should be taken as primitive conditions, which cannot be further analysed. I disagree with both views. It is important here to recall the distinction between cognitive and normative aims and explananda. One thing is to have a psychological state of being in joint attention. A different thing is to outline the justifications for that state. To aid maintain this distinction, I suggest differentiating between *psychological* joint attention and *normative* joint attention. The normative notion of joint attention consists of conditions (1) - (4). Conditions (3) and (4) themselves, however, refer strictly to the psychological states of A and B, states which need not include any concept, reflection or awareness as to the normative force they play when all conditions (1) - (4) are realised. In other words, the aetiology of these states need not involve the normative notion of joint attention. There is, therefore, no circularity, and there are no primitive unanalysable intersubjective states. Conditions (3) and (4), as psychological states, can and should be further analysed. The empirically based view proposed in the previous section is an attempt, after all, to sketch the cognitive capacities that go into the psychological state of pre-reflectively assuming that one is sharing attention to the same object with someone else.

Returning to the normative question, how are condition (1) - (4) realised? As noted above, the normative problem for A (as for B, *mutatis mutandis*) is not to *arrive* at the state that she is jointly attending to x with B. The problem A faces is to *avoid* defaulting to her prior assumption that she jointly attends to x with B, when that default should not be made — that is, in cases where B does not share her attention. The problem A faces is to avoid defaulting to her prior assumption that (3), when (3) should be rejected. There are two basic ways in which A's having a psychological state of assuming mutually shared attention to x with B is not rationally justified:

- i. Both conditions (2) and (4) don't hold.

- ii. Condition (2) holds but (4) doesn't. B attends to x, but does not have a psychological state of being attending to x together with A.⁵

Since A's psychological state of sharing attention to x with B does not need to include conditions (2) and (4), its rational justification is external to A's psychology. Her psychological state can be merely based on the following implicit reasoning: "I have a certain relation to x, and since you and I are so similar, (I assume) you have it too." Nothing requires A to reason any further. If she stops here, she takes herself to jointly attend to x with B (as a matter of *psychology*, that is). She may, of course, be epistemically wrong and unjustified. This is one possible stage of epistemic failure. Given (i) and (ii), to override defaulting to assumptions of mutual awareness, however, A will have to estimate the probability that B does not attend x, and the probability that B's attention is not being shared (see Siposova & Carpenter, 2019).

These estimates constitute two further psychological factors that modulate the rationality of her assumption of mutual awareness (or her withholding the assumption). On the proposed hypothesis, estimating these probabilities is cognitively demanding — though such process can of course (and likely it does) occur unconsciously and pre-reflectively. Depending on her prior experience in domains involving objects such as x and people such as B, A may be more or less sensitive to relevant information for estimating these probabilities. According to her estimates, then, A will reject her assumption of mutual awareness and conclude that there is no joint attention with B, or A will keep her assumption. The accuracy of her estimates, however, is a matter of degree, and cannot put A in a fully justified and rational state of *normative* joint attention, since these estimates will never fully encompass, from the psychological stance of A alone, the truths of conditions (2) and (4). These two conditions are external to A, and therefore beyond the ken of a strictly psychological standpoint. Paraphrasing Herbert Clark (1996: 96), a fully rational state of *normative* joint attention, i.e. one that is itself fully justified, consisting of conditions (1) – (4) above, can only be held by an omniscient being. The rationality of A's assumption of mutual awareness is, on this account, a matter of degree.

Does this mean that joint attention's epistemic justificatory role in our social lives is in jeopardy? Although the answer will ultimately depend on how strict we make the notion of epistemic rationality, I do not see compelling reasons to conclude that it doesn't. There is no a priori need to assume that the functional role of joint attention is to provide logically irrefutable grounds allowing individuals to engage rationally in joint endeavours

⁵ What about a situation where condition (4) holds but (2) doesn't? Whether this situation can ever occur will depend on whether (4) necessarily entails (2), which, in turn, will depend on the particular theory of perceptual attention endorsed. For example, one could erroneously assume oneself to be attending to x but be, in fact, attending to y. For simplicity, however, in this paper I assume that the entailment holds necessarily, so that the situation where condition (4) holds but (2) fails cannot occur.

concerning the jointly attended object. This normative assumption not only ignores psychological plausibility, but it may also leave aside aspects of strategic rationality at play in joint attentional scenarios (cf. Stein, 1996; Todd & Gigerenzer, 2012). Williams James already noted that “the logic of belief and knowledge” is too abstract to tackle many epistemic situations. We cannot simultaneously attempt to believe as many truths as possible and as few falsehoods as possible. According to James, we need to implicitly or explicitly weigh the value of avoiding false positives against false negatives (James, 1956; cf. Van Fraassen, 2002: 88).

In a football game, for example, if I assume joint attention with an opponent toward the ball and get it wrong, I may not incur any significant costs, given my practical interests (and I may never realise I got it wrong). But if I don’t assume joint attention, and should have (i.e. conditions (1), (2) and (4) hold, but A still fails to assume joint attention), the consequences could be drastic. Here, a false negative is more pernicious than a false positive. The coordinated attack scenario is an extreme case in the opposite direction: if I assume joint attention and get it wrong, then we both stand to incur high costs. But if I don’t assume joint attention, and should have, the costs are negligible. Here, a false positive is more pernicious than a false negative. It would seem rational for an agent to be sensitive to the pay-off structure of a particular situation. A notion of epistemic justification in joint attention based on strict objective accuracy, consisting of conditions (1) - (4), cannot capture such strategic aspects. A psychological state of being attending to an object or event together with another person may thus serve a functional role in action, without that state itself being necessarily accurate or irrefutably justified.

Conclusion

People effortlessly engage with others in activities that require attending together to some object or event. This ability of coordinated joint attention is considered to be fundamental to many aspects of human development, cognition, and interaction. It is widely held that joint attention is essentially public, or “out in the open”. Going beyond the metaphor of openness, however, requires a proper account of the mutual awareness that underlies joint attention. Current accounts, I have argued, fail to distinguish between two distinct explanatory aims when theorising on the openness of joint attention. One aim is normative: how should the openness of joint attention be characterised to account for its epistemic significance? Engagement in joint attention provides a rational basis for coordinated actions and shared knowledge about the world and others. A distinct explanatory aim is cognitive: what cognitive capacities and mental processes or understanding are involved in joint attention? Both reductive and non-reductive accounts of the openness in joint attention are primarily concerned with the normative aim, and their tensions arise when they extend their scope to address the cognitive aim.

Drawing from empirical research in infants and adults, I suggest that the openness in joint attention is not something that co-attenders must arrive at, but is implicitly assumed. On

this hypothesis, given the right sort of cues, people will tend to assume, often without any conscious reflection, that they are attending to some object or event together with someone else. Being able to entertain this assumption in the first place will require a set of minimal cognitive capacities, including the understanding that the other individual is a separate, live organism with their own goals, and that they engage with the world in a similar way to oneself. Ultimately, however, uncovering the set of cognitive capacities behind joint attention is an empirical project, not a purely conceptual one. Distinguishing between cognitive- and normative-first approaches to joint attention allow us to distinguish between two different notions of the concept. Joint attention as a *psychological* state requires an analysis at the level of the mental and psychological processing of the individual. Joint attention as a *normative* state — at the level of a “space of reasons”, and pertaining to justified rational epistemic states of the individuals — will have to include externalist conditions outside a person’s psychology. On this proposal, an individual cannot ever be fully aware of having rationally justified joint attention toward a common object with a third party, nor can they ever be fully aware of all the factors that make their state of joint attention rationally justified (cf. Sperber & Wilson, 1995: 19-20). They can, at best, form more or less accurate estimates of these factors. These estimates may be sufficient to account for the functional role of joint attention in supporting social behaviours and joint actions. If we are interested in providing a psychologically expedient construct of joint attention (Campbell, 2018; Eilan, 2005; Wilby, 2010), these considerations suggest that the notion of a fully normative state of openness in joint attention may well be cast aside.

References

- Akkoyunlu, E. A., Ekanadham, K., Huber, R. V., Akkoyunlu, E. A., Ekanadham, K., & Huber, R. V. (1975). Some constraints and tradeoffs in the design of network communications. *Proceedings of the Fifth ACM Symposium on Operating Systems Principles*, 9, 67–74.
- Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, 55(4), 1278–1289.
- Battich, L., & Geurts, B. (2020). Joint attention and perceptual experience. *Synthese*. doi: 10.1007/s11229-020-02602-6.
- Bermúdez, J. L. (1998). *The paradox of self-consciousness*. Cambridge, MA: MIT Press.
- Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, 18(5), 382–386.
- Brinck, I., & Gärdenfors, P. (2003). Co-operation and communication in apes and humans. *Mind and Language*, 18(5), 484–501.
- Bruner, J. S. (1995). From joint attention to the meeting of minds. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 1–14). Hillsdale, NJ: Lawrence Erlbaum.
- Calabi, C. (2008). Winks, sighs and smiles? Joint attention, common knowledge and ephemeral groups. In H. B. Schmid, K. Schulte-Ostermann, & N. Psarros (Eds.), *Concepts of sharedness: Essays on collective intentionality* (pp. 41–58). De Gruyter.

- Campbell, J. (2005). Joint attention and common knowledge. In N. M. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds. Issues in philosophy and psychology* (pp. 287–297). Oxford: Oxford University Press.
- Campbell, J. (2011). An object-dependent perspective on joint attention. In A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 415–430). Cambridge, MA: MIT Press.
- Campbell, J. (2018). Joint attention. In M. Jankovic & K. Ludwig (Eds.), *The Routledge handbook of collective intentionality* (pp. 115–129). New York, NY: Routledge.
- Carpenter, M., & Liebal, K. (2011). Joint attention, communication, and knowing together in infancy. In A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 159–182). Cambridge, MA: MIT Press.
- Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (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), i–174.
- Clark, H. H. (1992). *Arenas of language use*. Chicago, IL: The University of Chicago Press.
- Clark, H. H. (1996). *Using language* (pp. xi, 432–xi, 432). Cambridge: Cambridge University Press.
- Davidson, D. (1999). The emergence of thought. *Erkenntnis*, 51(1), 511–521.
- Eilan, N. (2005). Joint attention, communication, and mind. In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds. Issues in philosophy and psychology* (pp. 1–33). Oxford: Oxford University Press.
- Eilan, N. (2015). *Joint Attention and the Second Person (draft)*. <https://warwick.ac.uk/fac/soc/philosophy/people/eilan/jaspup.pdf>
- Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, 40(6), 760–768.
- Farrar, B. G., & Ostojić, L. (2018). Does social distance modulate adults' egocentric biases when reasoning about false beliefs? *PLOS ONE*, 13(6), e0198616.
- Fodor, J. A. (2003). *Hume variations*. Oxford: Clarendon Press.
- Gallagher, S. (2010). Movement and emotion in joint attention. In S. Flach, D. Margulies, & J. Söffner (Eds.), *Habitus in habitat I: Emotion and motion* (pp. 41–54). Bern: Peter Lang.
- Gallagher, S. (2011). Interactive coordination in joint attention. In A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 293–305). Cambridge, MA: MIT Press.
- Geurts, B. (2019). Communication as commitment sharing: Speech acts, implicatures, common ground. *Theoretical Linguistics*, 45(1-2), 1–30.
- Hobson, P., & Hobson, J. (2011). Joint attention or joint engagement? Insights from autism. In A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 115–136). Cambridge, MA: MIT Press.
- James, W. (1956). The will to believe. In *The will to believe and human immortality* (pp. 1–31). New York, NY: Dover Publications.

- León, F. (2021). Joint attention without recursive mindreading: On the role of second-person engagement. *Philosophical Psychology*, 34(4), 550–580.
- Lewis, D. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.
- Liddle, B., & Nettle, D. (2006). Higher-order theory of mind and social competence in school-age children. *Journal of Cultural and Evolutionary Psychology*, 4(3), 231–244.
- Mole, C. (2017). Attention. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/fall2017/entries/attention/>
- Moll, H., Carpenter, M., & Tomasello, M. (2011). Social engagement leads 2-year-olds to overestimate others' knowledge. *Infancy*, 16(3), 248–265.
- Moll, H., & Meltzoff, A. N. (2011). Joint attention as the fundamental basis of understanding perspectives. In A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 393–414). Cambridge, MA: MIT Press.
- Moore, C., & Dunham, P. J. (1995). Current themes in research of joint attention. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 15–28). Hillsdale, NJ: Lawrence Erlbaum.
- Mundy, P. (2018). A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder. *European Journal of Neuroscience*, 47(6), 497–514.
- Peacocke, C. (2005). Joint attention: Its nature, reflexivity, and relation to common knowledge. In N. M. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds. Issues in philosophy and psychology* (pp. 298–324). Oxford: Oxford University Press.
- Reddy, V. (2010). *How infants know minds*. Cambridge, MA: Harvard University Press.
- Reddy, V. (2015). Joining intentions in infancy. *Journal of Consciousness Studies*, 22(1-2), 24–44.
- Roessler, J. (2005). Joint attention and the problem of other minds. In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds. Issues in philosophy and psychology* (pp. 230–259). Oxford: Oxford University Press.
- Schiffer, S. R. (1972). *Meaning*. Oxford: Clarendon Press.
- Schmitz, M. (2015). Joint attention and understanding others. *Synthese Philosophica*, 29(2), 235–251.
- Seemann, A. (2019). *The shared world: Perceptual common knowledge, demonstrative communication, and social space*. Cambridge, MA: MIT Press.
- Siposova, B., & Carpenter, M. (2019). A new look at joint attention and common knowledge. *Cognition*, 189, 260–274.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Oxford: Blackwell.
- Stein, E. (1996). *Without good reason: The rationality debate in philosophy and cognitive science*. Oxford: Clarendon Press.
- Thomas, K. A., DeScioli, P., Haque, O. S., & Pinker, S. (2014). The psychology of coordination and common knowledge. *Journal of Personality and Social Psychology*, 107(4), 657–676.

- Todd, P. M., & Gigerenzer, G. (2012). What is ecological rationality? In *Ecological rationality: Intelligence in the world* (pp. 3–30). Oxford: Oxford University Press.
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Lawrence Erlbaum.
- Tomasello, M. (2003). *Constructing a language*. Cambridge, MA: Harvard University Press.
- Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.
- Tomasello, M. (2019). *Becoming human: A theory of ontogeny*. Cambridge, MA: Harvard University Press.
- Van Fraassen, B. C. (2002). *The empirical stance*. Princeton University Press.
- Wilby, M. (2010). The simplicity of mutual knowledge. *Philosophical Explorations*, 13(2), 83–100.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.
- Zawidzki, T. W. (2013). *Mindshaping: A new framework for understanding human social cognition*. Cambridge, MA: MIT Press.