



Joint attention and perceptual experience

Lucas Battich^{1,2}  · Bart Geurts^{3,4}

Received: 24 October 2018 / Accepted: 22 February 2020
© The Author(s) 2020

Abstract

Joint attention customarily refers to the coordinated focus of attention between two or more individuals on a common object or event, where it is mutually “open” to all attenders that they are so engaged. We identify two broad approaches to analyse joint attention, one in terms of cognitive notions like common knowledge and common awareness, and one according to which joint attention is fundamentally a primitive phenomenon of sensory experience. John Campbell’s relational theory is a prominent representative of the latter approach, and the main focus of this paper. We argue that Campbell’s theory is problematic for a variety of reasons, through which runs a common thread: most of the problems that the theory is faced with arise from the relational view of perception that he endorses, and, more generally, they suggest that perceptual experience is not sufficient for an analysis of joint attention.

Keywords Joint attention · Perceptual experience · Common knowledge · Relationalism · Perception · John Campbell

1 Introduction

Unbeknownst to each other, we are looking at the same piece of cake. Our attention is shared, but we don’t know that it is, and therefore the fact that our attention is shared doesn’t affect us in any way. But now we both come to realize that we are looking at the same piece of cake. Our visual attention becomes coordinated: alternating glances between the cake and the other, each is now aware of the other’s attention. Everything

✉ Lucas Battich
lucas.battich@campus.lmu.de

¹ Faculty of Philosophy, Ludwig-Maximilian-University Munich, Geschwister-Scholl-Platz 1, 80359 Munich, Germany

² Graduate School of Systemic Neurosciences, Ludwig-Maximilian-University Munich, Munich, Germany

³ Radboud University, Nijmegen, The Netherlands

⁴ HSE University, Moscow, Russian Federation

about our attention is out in the open, and if one of us were to say, “It’s mine!”, the intended referent would be obvious. Thus we went from a state of (merely) shared attention to joint attention. But what has changed? Or, as Hobson’s title (2005) has it, “What puts the jointness into joint attention?”

In recent years, the jointness question has received quite a lot of attention. It is agreed that joint attention is a ubiquitous phenomenon, and that it is important to human social interaction, because it helps us to coordinate our actions and beliefs. Since the term was introduced in developmental research by Jerome Bruner and colleagues (Bruner 1974; Scaife and Bruner 1975), joint attention has been considered a milestone in children’s social and cognitive development (Moore and Dunham 1995; Carpenter et al. 1998; Adamson et al. 2019), and shortcomings in joint attention have been associated with the onset of autism spectrum disorders (Hobson and Hobson 2011; Mundy 2016).

But there is no consensus on what joint attention *is*. The starting point for many (though not all) authors is that the jointness of joint attention is “open” between both attenders. It is fully and immediately transparent to them that they are jointly attending to the same object or state of affairs (thus, e.g., Bakeman and Adamson 1984; Tomasello 1995; Peacocke 2005; Calabi 2008; Campbell 2011; Carpenter and Liebal 2011; Eilan 2015). The challenge is to go beyond the metaphor of openness.

One view is that joint attention is to be understood in terms of common knowledge, or some related notion like common awareness, common belief, common acceptance, etc. There are various ways of fleshing out this idea, but the simplest is to define that we jointly attend to an object iff it is common knowledge between us that each of us is attending to it. Assuming that common knowledge can be defined in terms of knowledge (or a related epistemic notion like belief or awareness), this proposal is reductionist in the sense that it defines joint attention in terms of individual mental states.

An alternative to the knowledge-based approach is to view joint attention as a primitive relation, which is irreducible to the individual states of its relata (e.g., Calabi 2008; Seemann 2004). John Campbell’s (2005, 2011, 2018) “relational” theory is a prominent representative of this view. On Campbell’s account, when jointly attending to the cake, each of us *experiences* the other as jointly attending to the cake, such that you are a constituent of my visual experience, as I am a constituent of yours. Whatever epistemic significance joint attention has in coordinating our actions and beliefs, it results from this sensory experiential character. Unlike the knowledge-based approach, Campbell’s analysis is based on the idea that joint attention is “fundamentally a phenomenon of sensory experience” (2011, p. 415) and seeks to avoid referring to judgments, inferences, appeals to knowledge, beliefs, or any other higher cognitive processes. For this reason, Campbell’s style of analysis has proved to be attractive to cognitive and developmental psychologists (e.g., Moll and Meltzoff 2011; Hobson and Hobson 2011). Moreover, the relational view has been claimed to support and complement an interactionist approach to social cognition based on embodied, embedded, and extended interactive processes (Gallagher 2010; León et al. 2019). In these and other ways, the feasibility of the relational approach is critical to several lines of thinking in social cognitive psychology and the philosophy of mind and language.

We argue that Campbell’s relational account of joint attention fails to deliver on its promises. In doing so, we do not wish to advocate the knowledge-based approach, let

alone defend it against objections from Campbell or other critics. Rather, we intend to assess the merits of the relational approach in its own right, and will argue that, at several points, Campbell's theory threatens to collapse into its competitor. Therefore, we will need to discuss the knowledge-based view in some detail.

To begin with, Sect. 2 elaborates on the knowledge-based approach that serves as the foil to Campbell's relational account, which is presented in Sect. 3. Campbell's account is underdeveloped in several respects, and therefore we will need to consider various ways of making it more precise. We then argue that the relational definition of joint attention either results in an infinite regress of perceptual states or requires a construal of the notion of "co-attention" that is substantially identical with the notion of "normality" employed by knowledge-based theories, which, according to Campbell, shouldn't be part of an explanation of joint attention (Sect. 4). Finally, we discuss two further issues having to do with attention monitoring (Sect. 5) and failures of joint attention (Sect. 6).

The recurring theme throughout our discussion is that the problems which Campbell's theory runs into are due to tensions in his claim that joint attention is "fundamentally a phenomenon of sensory experience", and therefore not to be explained in terms of knowledge, belief, or awareness (2011, p. 415, 2018, p. 120). While joint attention undeniably involves sensory experience, our discussion suggests that an explanation of the phenomenon will have to factor in at least some knowledge, belief, or awareness.

2 The knowledge-based approach

Campbell presents his theory of joint attention as a superior alternative to the knowledge-based approach. On the latter view, joint attention can be defined as follows:¹

- A and B are jointly attending to x iff it is common knowledge between A and B that each of them is attending to x .

It doesn't matter for our purposes whether this particular definition is the best way of dealing with joint attention in terms of common knowledge; nor does it matter whether, e.g., common awareness, common belief, or common acceptance might be preferable to common knowledge. The only thing that matters is that all analyses that take this general approach have two features in common: they refer to cognitive states and they entail that these states give 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*.

Structures like this are the fingerprint of common knowledge, common belief, and so on (Lewis 1969; Schiffer 1972; Geurts 2019). On a knowledge-based account, it is this iterative structure that is held to capture the jointness in joint attention.

¹ Throughout this paper, we confine our attention to instances of joint attention and of common knowledge involving two participants.

It is important to be clear about the status of this iterative structure, for it is often misunderstood as implying that A and B cannot have common knowledge unless (i) they make an infinite number of inferences and (ii) they mentally represent the outcomes of all these inferences. The first misunderstanding was addressed by David Lewis (1969, p. 53) even before it arose: “Note that this is a chain of implications, not steps in anyone’s actual reasoning. Therefore there is nothing improper about its infinite length.” While this may not help very much to explain what common knowledge is, it should have sufficed to dispel the notion that it requires an infinite number of inferences as a precondition. Regrettably, Lewis’s remark was widely ignored in the subsequent literature, but that is as it may be: objection (i) is merely a denial of Lewis’s remark, and now that it has been dispensed with, objection (ii) falls by the wayside, too.

Having characterized common knowledge in terms of the iterative structure shown above, it remains to be seen how common knowledge is achieved. Following Lewis (1969) and Schiffer (1972), it is generally accepted that there are many types of finite situations, or “bases” as Lewis calls them, that generate common knowledge. People interacting with each other will soon find out that they share the same language, the same social background, the same hobbies, and so on, and any of these commonalities will serve as a basis for common knowledge. In the case of joint attention, part of the relevant basis will be that A and B take each other to know “that 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, p. 31). If this normality condition is not fulfilled, joint attention cannot be achieved. For example, if A knows that B’s low eye pressure seriously distorts her vision, then A and B cannot jointly attend to their cake.

The knowledge-based approach invites (but does not entail) the hypothesis that, apart from the normality condition, joint attention may be affected by practically any kind of background knowledge. Just as individual attention is driven by goals, intentions, and beliefs, so is joint attention. Tomasello (2014) notes that in joint attention I must be sensitive to the features of an object or situation that are relevant for you. Just following your line of gaze is not enough (cf. Moll and Tomasello 2007). To illustrate, consider the following scenario. If you point at a tree, I can follow your pointing gesture, but that does not tell me whether you are attending to the apples it carries, its smooth trunk, or the fungus on its bark. Perhaps we have been out foraging for apples or, alternatively, you have been tasked with the care of fungus-infected trees. My perceptual experience will be the same in both scenarios, since our lines of sight converge on the same tree. It is our shared background knowledge that enables me to determine which aspects of the visual scene you are focusing on, and to engage in joint attention with you.

The knowledge-based view is consistent with a range of positions on how much and what kinds of background knowledge are required for joint attention. However, since Campbell restricts his attention to knowledge-based theories that adopt the normality condition, we will make the same restriction here.

Campbell (2005) criticizes the knowledge-based approach for being cognitively demanding and psychologically unrealistic, because it requires infinitely many inferences and infinitely many levels of mental representation. The same objection is made

by Eilan (2005) and Carpenter and Liebal (2011), among others, and Calabi (2008) goes so far as to suggest that the theory entails that joint attention requires a grasp of the concept of infinity. As we have already seen, we maintain that this line of criticism is a spurious one, and we will consider it no further. Nevertheless, and regardless of how the knowledge-based approach fares, one can of course maintain that the relational theory presents a viable alternative. We now turn to that account.

3 The relational theory of joint attention

On an orthodox analysis, perceptual experiences are constituted not only by our surroundings, but also by our mental representations. Campbell's theory of joint attention extends his anti-representational theory of perception, which he calls "relational". On the relational view, "the phenomenal character of your experience, as you look around the room, is constituted by the actual layout of the room itself: which particular objects are there, their intrinsic properties, such as colour and shape, and how they are arranged in relation to one another and to you" (Campbell 2002, p. 116; see also Martin 2004; Travis 2004; Crane 2006). On this view, perception is not a matter of representing objects, but involves a non-representational relation between the perceiver and the token object perceived. Up to a point, this agrees with our intuitions. When we look around the room, we would normally say that we experience the room itself and its contents. Therefore, Campbell's account is a representative of what is sometimes called a "naïve realist" view on perceptual experience.

Campbell construes perception as a three-place relation: "S perceives x as being F", where the F-term stands for something like the aspect under which x is perceived. If I look around the room, I see the objects it contains as having certain intrinsic properties and being arranged in certain ways relative to one another and to myself.² Campbell doesn't discuss F-properties and -relations in any detail, which is unfortunate, because they play a key role in his account of joint attention, as we will see. However, for our purposes it suffices to note that, as defined by Campbell, perception is a purely extensional relation, which entails that, if F and G are the same, "S perceives x as being F" is equivalent to "S perceives x as being G".

Campbell considers perception to be a primitive relation, in the sense that it is not to be analyzed in such terms as "x causes S to have a representational content as of something being F", where S experiences this representational content (Campbell 2002, pp. 117-18). More generally, perception is a relation between subjects, objects, properties, and relations that is irreducible to other mental states.³

Campbell's analysis of joint attention builds on his relational account of perception:

On a relational view, joint attention is a primitive phenomenon of consciousness. Just as the object you see can be a constituent of your experience, so too it can

² This is Campbell's (2002) analysis. In later work, he introduces "standpoints" as a further ingredient of perception (2009). Although it may be that at least some F-relations can be subsumed under this notion, it seems to us that Campbell's views on "perceiving as" remain unaffected by its addition.

³ For critical assessments of the relational view of perception, see Burge (2005), Byrne and Logue (2008) and Nanay (2014). For attempts to reconcile the relational and representational views, see Nanay (2015) and De Sá Pereira (2016).

be a constituent of your experience that the other person is, with you, jointly attending to the object. This is not to say that in a case of joint attention, the other person will be an object of your attention. On the contrary, it is only the object that you are attending to. It is rather that, when there is another person with whom you are jointly attending to the thing, the existence of that other person enters into the individuation of your experience. The other person is there, as co-attender, in the periphery of your experience. (Campbell 2005, p. 288)

On this account, joint attention involves an object x and two individuals who experience each other as co-attending. That is to say:

- A and B are jointly attending to x iff
 - A perceives x as being co-attended by B, and
 - B perceives x as being co-attended by A.

Here the F-term of relational perception à la Campbell is instantiated with the property of being co-attended by the other. Thus, if I am alone eyeing the cake on the table, and you arrive to engage in joint attention with me, then there is a change in my perception of the cake: I now *see* it as being co-attended to by you.

But what does it mean for two people to co-attend to an object? Campbell doesn't say, but the most natural answer, it seems to us, is that co-attention is just a variant expression for joint attention: A and B perceive an object as being co-attended by the other iff they jointly attend to it. Alternatively, and perhaps less likely, co-attention and joint attention may be distinct concepts. We will consider both options shortly.

According to Campbell, joint attention differs from other forms of simultaneous attention in two respects. If A and B are jointly attending to x , then first, A and B monitor each other's attention, and second, A's attention is one of the factors controlling B's attention, and vice versa (cf. Tomasello 1995, p. 107). In line with his relational principles, Campbell sees attention monitoring and control as relations that are to be fleshed out in causal terms. For example, Campbell stipulates that, in order for joint attention to be achieved, B's continued attention to x must be one of the causal factors for A's continuing to attend to x , and vice versa, A's continued attention to x must be one of the causal factors for B's continuing to attend to x (Campbell 2005, p. 289).

Campbell assumes that "this coordination of attention may involve the use of sub-personal mechanisms, rather than explicit, personal-level thoughts about the direction of the other person's attention" (2005, p. 288). Here the personal/subpersonal distinction coincides with the introspectable/non-introspectable distinction, and personal mental states are taken to include sensations, emotions, beliefs, desires, and rational and deliberate thinking. The subpersonal processes for monitoring and control cannot account for the jointness of joint attention, because joint attention is a personal-level phenomenon and "it is hard to see what [these subpersonal processes] contribute to the subject's psychological life" (Campbell 2011, p. 416). The jointness in joint attention is a personal-level state. This is why Campbell seeks to explain joint attention in terms of the conscious experience of having the other as a co-attender. On Campbell's view, it is precisely in virtue of its perceptual experiential character that joint attention can have the epistemic significance that it has.

Campbell maintains that his relational theory of joint attention is free of the difficulties that he and others associate with the knowledge-based approach. First, it doesn't involve the infinitely iterating structures that are the hallmark of knowledge-based theories. Second, it doesn't appeal to background knowledge, and in particular, it doesn't appeal to anything like the normality condition, which on a knowledge-based account is instrumental in generating these iterative structures. In short, on a relational analysis, joint attention is defined in terms of perceptual experience (Campbell 2011, p. 415). In the following sections, we argue that this position is untenable.

4 Co-attention

As a matter of logical necessity, co-attention and joint attention are either the same thing or not. Peacocke has noted that, if co-attention and joint attention are the same thing, the notion of a co-attender presupposes the property which is to be explained, i.e. the openness of joint attention (2005, p. 300). Nevertheless, Campbell's own discussion suggests rather strongly that, for him, joint attention and co-attention are identical: to be a co-attender is just to stand in the primitive three-place experiential relation, with another co-attender, to a common object (2011, p. 420). Since joint attention is an extensional relation, this entails that, if $x = x'$, then A and B are jointly attending to x if and only if they are jointly attending to x' (2011, p. 424). Thus, it follows that, if A and B jointly attend to x:

- A perceives x as being co-attended by B,
- B perceives x as being co-attended by A,
- A perceives x as being perceived by B as being co-attended by A,
- B perceives x as being perceived by A as being co-attended by B,
- A perceives x as being perceived by B as being perceived by A as being co-attended by B,
- B perceives x as being perceived by A as being perceived by B as being co-attended by A,
- and so on *ad infinitum*.

So now joint attention involves the same sort of infinite iterations that, according to Campbell, invalidate the knowledge-based approach. Hence, in this respect, Campbell's theory turns out to mimic the knowledge-based approach. By his own lights, this is an unwanted result, since part of the motivation for claiming that "joint attention is a primitive phenomenon of consciousness", is to avoid the complex iterations of mental states that plague the knowledge-based approach (Campbell 2005). While Campbell is not fully clear on the notion of primitiveness, it is often assumed that an infinite regress is blocked just because joint attention is a primitive relation (e.g. Calabi 2008; Seemann 2004; Eilan 2015). We fail to see how this line of defence might work. Consider the following case. In most versions of propositional logic, conjunction is a primitive, non-reductive relation, in the sense that it cannot be reduced to other relations (other versions may take disjunction or implication to be primitive instead). This doesn't prevent it from licensing endless series of inferences. For example, if $p \ \& \ q$ holds, then:

- $p \ \& \ p \ \& \ q$,
- $p \ \& \ p \ \& \ p \ \& \ q$,
- $p \ \& \ p \ \& \ p \ \& \ p \ \& \ q$,
- and so on *ad infinitum*.

Evidently, primitiveness in itself does not impede recursion, and there is no reason to suppose that the primitiveness of joint attention will block the infinite regress pictured above. As Campbell's defines it, co-attendance may not be reducible to other individual mental states, but that doesn't prevent the definition from generating a recursive regress. What is crucial to the regress is that the relation is extensional. If primitiveness is meant to resolve the issue, we are owed a positive explanation of what it is and how it accomplishes this feat.

Thus, the assumption that co-attention is joint attention, which Campbell seems to subscribe to, leads into major trouble for his account. Therefore, let's consider the possibility that co-attention and joint attention are not the same thing, and let's grant, if only for the sake of the argument, that this will block the infinite regress that would otherwise ensue. According to Campbell, when we are engaged in joint attention there is a difference between how I am related to the co-attended object and how I am related to you. Each person is "there" and enters the other's experience, "as co-attender" (Campbell 2011, p. 419). We will not try to provide a full-dress definition of co-attention, but will merely consider what are likely to be some of the minimal conditions that must hold for someone to enter another person's perceptual experience as a co-attender.

In order to experience B as co-attender, A must be able to recognize that B co-attends to x with her. Apart from the fact that this is a natural assumption to make, it is also in line with what Campbell writes about other F-properties:

To experience the shape of a solid object you must have some capacity to recognize manifest sameness of shape across movements by you or by the object. Otherwise it is hard to see how you could be said to be encountering the property of three-dimensional shape at all. (Campbell 2009, p. 288)

By analogy, having the capacity to recognize co-attention when one encounters it is a necessary precondition for joint attention. What does this capacity involve? For starters, a plausible candidate is the ability to recognize the other as an animate entity, separate from oneself, and to sense, however minimally, the other's agency (e.g. that they have goals different from one's own). But clearly this won't suffice for me to see you as a co-attender rather than merely as a person who happens to be looking in the same direction, for example, or who is incapable of looking in the first place. Thus we are led to suppose that the ability to recognize co-attention requires the ability to include as candidate co-attenders people whose line of sight intersects with the target object, and exclude the blind, blindfolded, comatose, and so on. But these are precisely the sort of requirements that make up the normality condition on which the knowledge-based view is based.

This is bad news for Campbell's account for two reasons. First, because he explicitly seeks to avoid any appeals to knowledge, beliefs, or awareness of the two participants (Campbell 2018, p. 120). Secondly, because on the knowledge-based view, the nor-

mality condition is the linchpin in generating the endless iterations of psychological states that Campbell rejects. Again, we have come to a point at which the relational account threatens to converge with the knowledge-based account.

The key observation on which the foregoing argument is based is just that the kind of *content* that seems to be needed to flesh out the notion of co-attention is the same as what goes into the normality condition. No assumptions have been made about the *nature* of that content, except for the fact that, if this content is a prerequisite for my joint attentional perceptual experience, it cannot itself be accounted for in terms of that experience. This point bears emphasizing because the normality condition has been held to require a grasp of a concept of psychological normality and all that goes with it (e.g., Peacocke 2005). As far as we can tell, that is not the case. Whatever kind of content is involved in co-attention will work for normality, too (cf. Calabi 2008).

As discussed in Sect. 2, while the knowledge-based view is consistent with the hypothesis that, in principle, joint attention may be affected by any kind of knowledge, it allows for a range of positions on how much and what kinds of knowledge are required for joint attention; the normality condition may be seen as an attempt at capturing at least some of that knowledge. In the foregoing we were led to conclude that, whatever co-attention may be, it seems likely to involve the same kind of knowledge. The bottom line is that *at least some* knowledge must be involved in any analysis joint attention. If we try to do without any form of knowledge whatsoever, a feasible account of joint attention is outside our reach. Therefore, the relational view is on the wrong track. In the remainder of this paper we discuss two further issues that reinforce this conclusion.

5 Causal monitoring

On Campbell's view, when we are engaged in joint attention, you are a constitutive part of my experience. For this to happen, some causal conditions must be met (Campbell 2005, p. 288). Part of your causal contribution to my experience is that you are continuously attending to the object that I'm attending to. More formally:

- A's continued attention to *x* must be one of the factors causally sustaining B's continuing to attend to *x*, and
- B's continued attention to *x* must be one of the factors causally sustaining A's continuing to attend to *x*.

These conditions can be interpreted strongly or weakly. On a strong interpretation, causal monitoring must be literally continuous, i.e. uninterrupted. This interpretation is suggested by Campbell's (2005, p. 289) own words, which we have reproduced almost verbatim. On the weak interpretation, monitoring need not be continuous in order to sustain joint attention. It is not hard to see that, in both versions, Campbell's causal conception of monitoring is problematic: on the strong interpretation it is unrealistic, and on the weak interpretation it is hard to see how it could be causal. Since the two interpretations are jointly exhaustive, it is doubtful that the mutual monitoring on which joint attention is generally agreed to be based is a causal relation.

Consider the strong version first. On this interpretation, you have to keep looking at our cake without interruption in order for our state of joint attention to persist. If you divert your gaze even for a second, the causal connection is broken, I cease to experience you as a co-attender, and our joint attention is no more. This is clearly wrong. When we jointly attend to our cake, for example, we typically alternate gazes between the cake and each other; if both of us were fixedly staring at the cake without checking with each other every once in a while, we wouldn't be engaged in joint attention. Hence, on a strong interpretation, causal monitoring fails to account for the facts.

On a weak interpretation, your eye gaze is allowed to shift between the cake and myself (and perhaps other objects as well). But then how can we account in purely causal terms for the difference between a solid 10-minute bout of joint attention and an interval of the same length during which our joint attention is briefly interrupted every now and then? On a knowledge-based model, joint attention may be sustained, in part, by informational processes that enable us to distinguish between genuine interruptions and merely apparent ones. For example, if it is common knowledge between A and B that each is equally interested in *x* and the other, then alternating gazes between each other and *x* are more likely to be experienced as joint attention than if it is common ground that *x* is of predominant interest for both A and B. It is hard to see how such observations could be accommodated by a purely causal model and the perceptual relation it is meant to support. To explore this point a bit further, we turn to our last topic: failures of joint attention.

6 When joint attention fails

Once again, we have been jointly attending to our cake for a while, when you start daydreaming about your next holiday, and although you're still gazing at the cake, your mind is now elsewhere. Hence, our episode of joint attention has come to an end, but as far as I'm concerned we are still looking at the cake together. How is this possible? In his 2011 article Campbell diagnoses the situation as follows:

Being an experiential relation, like “___ sees ___”, it is introspectable: X can tell just by reflection that he or she is co-attending with Y to Z. However, here as so often, introspection is not an infallible source of knowledge. You may think you are co-attending with Y to Z even though Y left long ago. (Campbell 2011, p. 419)

Based on introspection, I believe that we are jointly attending to the cake while in fact we're doing no such thing anymore. On Campbell's account, this is a scenario that theories of direct perceptual experience are all too familiar with. Consider the following experiment. A subject is looking at a tennis ball which, during the 200 milliseconds of an eye blink, is replaced with another, qualitatively indistinguishable ball. So our subject doesn't notice the change, and as far as she is concerned her perceptual experience is the same as before. For a relational theorist like Campbell the case is clear cut: the replacement causes the subject to enter a new perceptual state, even if she fails to notice it (cf. Martin 2004; Schellenberg 2010).

This claim is controversial, but it is the logical consequence of the premise that, in veridical perception, external objects and their properties “partly constitute one’s conscious experience” (Martin 1997, p. 83). This premise clashes with the intuition that two conscious perceptual experiences that are indistinguishable for a subject are necessarily the same (Martin 1997, p. 81). It is generally agreed that these two notions are difficult if not impossible to square. However, we will not address that issue here, and merely want to point out that, compared to the tennis ball experiment, cases of false joint attention raise additional issues for the relational account.

First, whereas in the tennis ball experiment the perceived objects are numerically distinct, in our case of failing joint attention it is just the fact that you cease to pay attention to the cake that, according to Campbell, causes a change in my perceptual experience. By hypothesis, there are no external factors that might causally account for my change of perceptual state. Only *neural* changes in you might conceivably qualify for this job. Therefore, Campbell owes us an account of how covert changes in the brain states of one person can affect perceptual experiences in another.

Secondly, since the relational view allows for dissociations between my perceptual experience and my beliefs about my perceptual experience, it also allows for the possibility that I am engaged in joint attention but mistakenly believe that I am not. But this seems to be at odds with the key feature of joint attention that, as noted in the introduction, all parties agree on: joint attention is public, it has a special kind of openness or mutual manifestness. On Campbell’s account, this openness is constituted by my perceptual experience and yours, and therefore I can mistakenly believe that we are not engaged in joint attention because I am wrong about my experience. This sounds like a downright contradiction to us. It is one thing to suppose that I can wrongly believe that we *are* engaged in joint attention; this is a possibility that every theory should allow for. But it is quite another thing to suppose that I can wrongly believe that we *are not* engaged in joint attention. This is a possibility that, in our view, should be ruled out by the very notion of joint attention, and if this much is true, it is problematic that Campbell’s account fails to do so.

7 Conclusion

The relational approach has been touted as a superior alternative to the knowledge-based approach, and has been claimed to provide an account of joint attention anchored in its perceptual experiential character, which avoids an infinite regress of inferences, does not require conceptual understanding, and generally imposes minimal demands on processing and representation. For these reasons, it has proved to be attractive to philosophers and psychologists concerned with social cognition and its development.

In the foregoing we have argued that Campbell’s theory is untenable, and at several points comes perilously close to collapsing into its knowledge-based competitor. To begin with, the relational analysis either results in an infinite regress of psychological states or necessitates elaborations of the notion of “co-attention” that make it indistinguishable from the notion of “normality” employed by knowledge-based theories. Further, the theory requires a causal notion of attention monitoring which is either too strict to be realistic or so loose that it cannot be a purely causal notion in the

first place. Finally, the theory implies a counter-intuitive dissociation between my perceptual experience and my beliefs about my perceptual experience, which becomes problematic when considering cases of joint-attention failure.

Where do we go from here? The relational view is predicated on the assumption that joint attention is fundamentally a type of primitive perceptual state, not itself susceptible to explanation in terms of the knowledge, beliefs, or awareness of the two participants (Campbell 2018, p. 120). Our discussion suggests that the difficulties the relational view faces may be tackled by abandoning this assumption, and by taking into account the knowledge, beliefs, or informational states of each participant in joint attention. Of course, joint attention certainly includes perceptual experiential aspects. Our discussion of the relational view suggests, however, that these perceptual aspects are not sufficient for an account of joint attention. Further work in a theoretical explanation of joint attention may do well by taking into account the combination of sensory experience and individual epistemic states.

To sum up, we believe that our arguments raise serious issues for theories of joint attention that adopt a relational take on perceptual experience. More generally, they suggest that theories anchored in perceptual experience will have to factor in at least some knowledge, belief, or awareness into the analysis of joint attention.

Acknowledgements Open Access funding provided by Projekt DEAL. We would like to thank the reviewers of *Synthese* for their very extensive comments. Bart Geurts was supported by the HSE University Basic Research Program.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adamson, L. B., Bakeman, R., Suma, K., & Robins, D. L. (2019). An expanded view of joint attention: Skill, engagement, and language in typical development and autism. *Child Development, 90*(1), e1–e18.
- Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development, 55*(4), 1278–89.
- Bruner, J. S. (1974). From communication to language: A psychological perspective. *Cognition, 3*(3), 255–287.
- Burge, T. (2005). Disjunctivism and perceptual psychology. *Philosophical Topics, 33*(1), 1–78.
- Byrne, A., & Logue, H. (2008). Either/or. In A. Haddock & F. Macpherson (Eds.), *Disjunctivism: Perception, action, knowledge* (pp. 314–19). Oxford: Oxford University Press.

- 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). Frankfurt: De Gruyter.
- Campbell, J. (2002). *Reference and consciousness*. Oxford: Oxford University Press.
- 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* (pp. 287–297). Oxford: Oxford University Press.
- Campbell, J. (2009). Consciousness and reference. In A. Beckermann, B. P. McLaughlin, & S. Walter (Eds.), *The Oxford handbook of philosophy of mind* (pp. 648–662). 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. 415–30). 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: 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.
- Crane, T. (2006). Is there a perceptual relation? In T. S. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 126–146). Oxford: Oxford University Press.
- De Sá Pereira, R. H. (2016). Combining the representational and the relational view. *Philosophical Studies*, 173(12), 3255–3269.
- Eilan, N. (2005). Joint attention, communication, and mind. In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds* (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>. Accessed October 9, 2018.
- Gallagher, S. (2010). Joint attention, joint action, and participatory sense making. *Revue de Phénoménologie*, 18, 111–124.
- Geurts, B. (2019). Communication as commitment sharing: Speech acts, implicatures, common ground. *Theoretical Linguistics*, 45, 1–30.
- Hobson, P. (2005). What puts the jointness into joint attention? In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds* (pp. 185–204). Oxford: Oxford University Press.
- 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.
- León, F., Szanto, T., & Zahavi, D. (2019). Emotional sharing and the extended mind. *Synthese*, 196(12), 4847–4867.
- Lewis, D. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.
- Martin, M. G. F. (1997). The reality of appearances. In M. Sainsbury (Ed.), *Thought and ontology* (pp. 81–106). Milan: Franco Angeli.
- Martin, M. G. F. (2004). The limits of self-awareness. *Philosophical Studies*, 120(1–3), 37–89.
- 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.
- Moll, H., & Tomasello, M. (2007). Cooperation and human cognition: The Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480), 639–648.
- 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. C. (2016). *Autism and joint attention: Development, neuroscience, and clinical fundamentals*. New York: Guilford Publications.
- Nanay, B. (2014). Empirical problems with anti-representationalism. In B. Brogaard (Ed.), *Does perception have content?* (pp. 39–50). Oxford: Oxford University Press.

- Nanay, B. (2015). The representationalism versus relationalism debate: Explanatory contextualism about perception. *European Journal of Philosophy*, 23(2), 321–336.
- 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* (pp. 298–324). Oxford: Oxford University Press.
- Scaife, M., & Bruner, J. (1975). The capacity for joint visual attention in the infant. *Nature*, 253, 265–266.
- Schellenberg, S. (2010). The particularity and phenomenology of perceptual experience. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 149(1), 19–48.
- Schiffer, S. R. (1972). *Meaning*. Oxford: Clarendon Press.
- Seemann, A. (2011). *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience*. Cambridge, MA: MIT 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. (2014). *A natural history of human thinking*. Cambridge, MA: Harvard University Press.
- Travis, C. (2004). The silence of the senses. *Mind*, 113(449), 57–94.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.