

From Joint Attention to Common Knowledge

Michael Wilby

Anglia Ruskin University, Cambridge

The Shared World: Perceptual Common Knowledge, Demonstrative Communication, and Shared Social Space. Axel Seemann. Cambridge, Massachusetts: MIT Press, 2019, 248 pages, \$40.00 hardback.

What is the relation between joint attention and common knowledge? On the one hand, the relation seems tight: the easiest and most reliable way of knowing something in common with another is for you and that other to be attentively aware of what you are together experiencing. On the other hand, they couldn't seem further apart: joint attention is a mere perceptual phenomena that infants are capable of engaging in from nine months of age, whereas common knowledge is a cognitive phenomenon involving (so it seems) complex, overlapping metarepresentational states that require the kind of sophisticated mindreading skills that developmental psychology has shown to be beyond the capabilities of young children.

In *The Shared World: Perceptual Common Knowledge, Demonstrative Communication, and Shared Social Space*, Axel Seemann attempts, inter alia, to make sense of this conundrum. The general answer that Seemann provides is what might be called a “bottom-up” one: joint attention, on Seemann's account, is understood in a minimalist way, such that it occurs via sub-personal and sub-contentful resonances between agents engaged in joint activity within a shared social environment. This allows for a later-developing reflective and conceptualised understanding of those joint engagements, including the capacity for full-blown perceptual common knowledge, with its attendant metarepresentational mindreading skills.

There are a couple of significant advantages for this approach. First, as Seemann demonstrates in the final chapter, it remains in line with the empirical evidence on the development of infants' and young children's capacities for understanding social interactions, which appear to follow something like the trajectory that Seemann's theory paints. And, second, it allows for a grounding for joint perceptual experience in something more homely than the kind of metaphysical and mathematical exotica of set theory, epistemic logic, game theory and possible world semantics in which puzzles of common knowledge usually reside, which in turn helps bridge the gap between joint attention and common knowledge.

As befits the subject matter, the book crosses a range of different areas, including philosophy of language, philosophy of perception, philosophy of mind, philosophy of action, epistemology, as well as engaging with current thinking in developmental psychology and social neuroscience. It is well written, full of provocative suggestions and arguments, with a seemingly never-ending supply of imaginative thought-experiments. It is a welcome addition to the literature that is likely to help shape future thinking on joint attention. As editor of one of the key interdisciplinary volumes on joint attention, Seemann (2011a) is clearly very well placed to bring together all this material and to try to make sense of it.

This review will follow the four-stage structure of Seemann's book (Perceptual Common Knowledge; Communication; Shared Social Spaces; Joint Attention). The aim shall be to provide a charitable, but critical, reconstruction of Seemann's overarching argument and then to assess its success at relating common knowledge and joint attention.

Perceptual Common Knowledge

The first section of the book is essentially an exercise in social epistemology. Seemann's concern here is to establish what would be required for "the facts to be out in the open" between two or more subjects (p. 16). This is what is sometimes referred to as *common knowledge*.

What is common knowledge? The general idea is that certain facts are common knowledge when all subjects within a group know the facts in question and know that they all so know (we will flesh this out in more detail shortly). Seemann's focus is on a particular variant of common knowledge: perceptual common knowledge (PCK). He wants to know how situations in which two participants share a perceptual awareness of their environment can allow for demonstrative communication. Seemann rejects the standard "iterative approach" to common knowledge (which he calls "a conjunctive analysis"). The iterative approach has it that the full openness of common knowledge can only be captured via an infinite series of overlapping mental states of the following form:

1. S knows that p
A knows that p
2. S knows that A knows that p
A knows that S knows that p
3. S knows that A knows that S knows that p
A knows that S knows that A knows that p
...etc.
...etc.
...etc. ad infinitum.

Seemann rejects this partially on the grounds that it runs into Gettier-style problems. Although his argument for this particular claim is (to my mind) rather opaque (pp. 20–21), the key point he wants to make is clear:

[I]nternalism ... is unattractive for the characterisation of perceptual common knowledge: its social character [can] only be captured in an account of the relation between individuals' beliefs, and such an account is vulnerable to the threat posed by the regress of infinite layers of knowledge that are generated by what each joint perceiver knows. (p. 21)

In its stead, Seemann argues for an “externalism” with regards to PCK. There are a number of notions, borrowed from Williamson (2000), that Seemann uses to elucidate this idea (“transparency” “luminosity” “factivity”), but central to the proposal is what Seemann calls a “joint perceptual constellation”:¹

If we are jointly perceiving an F, the perceptual constellation between you, me and the F is constituted by the respective locations of you, me and the F. So, you can think of perceptual constellations as kind of spatial relations that inform the way that the object is presented to its perceivers: the spatial relation between perceiver and object has an impact on the character of the perceiver's experience of the thing.... [T]he spatial relation changes fundamentally if the relation is triadic: if there are two individuals who jointly perceive an object, the spatial relation must be thought of as obtaining between the three constituents of the triangle. (p. 3, ft 7)

As I understand it, the idea is that joint perceptual constellations — i.e., the triadic spatial relations that hold between perceivers and what is perceived — help avoid the regress that Seemann argues is inherent in internalist characterisations of PCK. Whereas internalist characterisations of PCK are forced to trigger a set of never-ending overlapping embedded mental states, the externalist alternative can offload this burden into the external environment, where there is a ready-made triadic spatial relation constituted by the embodied perceiving subjects and the perceived object. Does this resolve the epistemological problem of PCK? Seemann thinks so:

¹Curiously, for such a central concept, its characterisation (which I am quoting) is relegated to a footnote.

... the spatial order of the environment individuates each perceiver's factive mental states of knowing the location of the object in social space, and it guarantees that the proposition expressing the location of the object in social space is commonly known.... [J]oint perceivers do not need to generate $(n+1)$ iterations of what each of them knows in order to obtain primary perceptual common knowledge: what each individual knows (the object's location in social space) is determined by what they know in common.... The joint constellation contains the regress. (p. 79)

There are two worries that one might have about this approach, I think. The first is that one might think that the approach does not do enough to explain how the regress is prevented. If there is a joint perceptual constellation that is constituted by the extensional spatial relations that hold between the triad, do we not still need to hypothesise that the agents are *aware* of this constellation holding? And, if that is the case, are we not then back on the same regressive road? For, if the joint perceptual constellation is merely a set of spatial relations, then, if taken on its own, it is too lean to provide the epistemological clout required to solve, say, a problem of coordinated attack:² spatial relations, in and of themselves, have no cognitive or epistemological power at all; if they are to do anything, then they need to be the content — perhaps nonconceptual — of a mental attitude of some kind. On the other hand, if we do implant them into a mental attitude, such that it is the agents' *awareness* of that underlying triadic relation that is playing the epistemological role, then this seems to just set off the regress again (for each agent now needs to know that the other is aware...and so on).

The second concern is more general: I have been following Seemann in thinking of the regress as somehow foisted upon the internalist. But this perhaps underplays its role in the internalist's thinking. The invocation of an infinite series of overlapping mental states primarily arises not because the internalist has no other way of being able to characterise the intuitive phenomenology of PCK — for perhaps a few iterations would suffice for this — but because, by the internalist's lights, neither rational joint intentional action nor communication would be achievable without it. That is to say: the regress that Seemann (rightly) finds objectionable in the internalist's account is itself intended as a solution to the question of how agents might be able to rationally coordinate their actions. That problem is one in which agents need to act together if they are to act at all, a problem in which, as Margaret Gilbert puts it, "you only have reason to do your own part...if you have reason to think the other player will" (Gilbert, 1989, p. 335), and vice versa. It is this looping rationalisation for joint intentional action (you have reason to act only if I do, and I have reason to act only if you do) that is problematic for internalist/individualistic approaches to coordination, and that forces their hand towards the iterative approach to common knowledge. What this all means is that although there is a potential solution to coordination problems that

²For a clear discussion of the problem of coordinated attack (aka "The General's Paradox") see Syverson (2003)

can be had by appealing to joint perceptual constellations and social spaces, such a solution needs to do two things: it must be rich enough to have the epistemic resources to justify rational joint action, yet also lean enough to avoid requiring overlapping iterations of mental states.

Demonstrative Reference and Communication

Although I have talked above largely about the role of PCK within rational joint actions and coordination, the bulk of Seemann's book focuses on the role of PCK within demonstrative communication. Part of the motivation for this is that Seemann thinks that PCK is by its nature communicative:

Imagine two strangers who, standing uncomfortably close to each other, are focusing on the same picture in a crowded gallery. Each of them is keenly aware of the other and each is at pains not to let the other interfere with her viewing experience: there is no pointing, no verbalising, no intentional gaze following, no exchange of glances. It is not credible that the two spectators should possess common knowledge about what they see....For *p* to be commonly known, it is necessary that each subject know that each subject knows that *p*. Meeting this demand requires them to communicate. (p. 36)

This claim is quite compelling on its face. What enables the agents in the example to know that they are focussing on the same thing is that their awareness is communicated.³ For how else is the shared information — as well as the fact that it is shared — to get from one subject to another? And, for Seemann, PCK is “typically acquired by way of linguistic communications...that contain demonstrative expressions” (p. 35). It is face-to-face communication about “this” and “that” — those things in the immediate environment that can be pointed out or gestured towards — that lead to PCK. The arguments in this section are at times dauntingly dense and complex, drawing on Frege, Evans, Kaplan, and Grice among many others. There are three claims that I think worth highlighting and from which we can summarise the core of Seemann's argument.

First, Seemann takes a broadly Gricean line on communication, identifying the content of a demonstrative utterance with the speaker's informative intention. *Successful* communication is then identified with the hearer's recognition of that intention (p. 55). PCK then explains *how* the hearer is able to identify the speaker's intention (the hearer can commonly identify the referent of the demonstrative utterance) and thus how successful communication is possible:

In order for a speaker to be in a position to inform a hearer of a perceptual fact about a jointly perceived object by means of a demonstrative utterance, speaker and hearer must perceptually know in common which object the speaker is intending to refer to and communicate about. (p. 38)

³ See also Davidson (2001, p. 121).

At first glance, one might worry that the account seems circular (for haven't we now said that PCK is both a necessary precondition for communication but also acquired via communication?). But the key motivating thought here, I assume, is this: PCK simply *is* successful (demonstrative) communication. If one has successfully communicated the thought that <this (gesture) is a hand> to another person, then one has established PCK that there is a hand here. And if one has established PCK that there is a hand here, then that can only be because one of us has successfully communicated the thought that <this (gesture) is a hand>. This, I think, makes sense of the gallery example quoted above.

Second, Seemann draws a distinction between a case in which one is *establishing* PCK towards an object, and a case in which one can *extend* one's PCK of an object by pointing out various features of it. This second type of case nicely illustrates the potential epistemic power of PCK. To adapt one of Seemann's examples: suppose that we are both looking at a Turner painting in a gallery, and there is an orange mark on the painting. Perhaps you know this:

1. *that (the mark) is a representation of a buoy.

And I know this:

2. *that (the mark) was a last-minute addition to the painting.

In communicating this information we can now extend our shared knowledge:

3. *that (the mark) is a representation of a buoy and was a last-minute addition to the painting.

What allows for the transition from (1) and (2) to (3) is the fact that we are speaking of the same object. But, note, that the identification of the object is entirely contextual; we are not relying on a description of the painting that we could each use and then attach the predicates to. Rather, we are relying on the demonstratively identified object: *that thing there*. And, as Seemann emphasises, this would not be available unless the interlocuters were sharing their perceptual space. Indeed: suppose there are high stakes involved. For instance, suppose that the following week you are a contestant on *Who Wants to Be a Millionaire?* and you have reached the final crunch question, which is: "Which Turner painting includes a representation of a buoy which was also a last-minute addition to the painting?" Should you (a) gamble, and answer the question, winning £1 million if correct, and lose all the money if not. Or should you (b) not answer the question, and just take the half million on offer on the table in front of you. Seemann's thought here is that you are justified in answering the question if you had PCK of the orange mark during the conversation. If, on the other hand, you had each simply said your part, without awareness of what each other was referring to, then the transition from (1) and (2) to (3) would be unjustified, since we wouldn't have been aware that we were discussing the same thing.

The third point to focus on from this section is this: if demonstrative identification within joint perceptual contexts is to take on the immediate, non-inferential role of expanding a group's knowledge, by allowing for discussion about the demonstratively identified object in the absence of an informative and communicated assertion of identity, then it would seem to require that the agents are not just referring to the same object in extension, but are also doing so in intension. At least it might seem that way. But Seemann appears to demur, and I think that examining why this is so will get us to the heart of his proposal. Seemann thinks that co-reference will suffice for a shared understanding of the content of a demonstrative:

Fregean senses are individuated in terms of their cognitive significance for individual subjects at particular times. But in certain situations, two subjects can immediately extend their knowledge about an object by considering arguments whose premises contain singular terms that co-refer to the same object, despite associating them with different cognitive significances. (p. 48)

Now, one might be concerned that this leaves the proposal open to a familiar family of counterexamples:⁴ for instance, we might know in common (because you say so while pointing) that <*that is a representation of a buoy>; and we might also know in common (because I say so while pointing to the same mark) that <*that is a last-minute addition to the painting>. But, for all that, we might still not be justified in, or capable of, drawing the shared conclusion that <*that is a representation of a buoy and a last-minute addition to the painting>, if one or other or both of us were to have lost sight of the fact that *that is identical with *that.

For this reason, one might think that an appeal to Fregean sense is required after all: if we share the same sense of *that while communicating (if we perceive it under the same “mode of presentation”), then we can simply trade on the identity that that shared mode of presentation provides. However, Seemann argues that an appeal to (Fregean) sense would be unnecessary. It is unnecessary because, he thinks, we have a joint perceptual constellation — a shared spatial framework — to do the work for us. It can do this because the interlocuters can co-identify *that as occurring at a uniquely identifying place within a shared spatial framework. We commonly know that we are referring to the same object because we are operating within a joint perceptual constellation.⁵

To illustrate the idea, imagine that this shared perceptual space is marked with a grid, with each square of the grid numbered from (say) 1–100. Given that the

⁴ See, for instance Greg McCulloch's (1984, pp. 199–208) discussion of the Passengergobbler.

⁵ What gets a little bit confusing in text is that Seemann at times seems to endorse a shared sense (see, e.g., p. 48). I think, however, that Seemann is (following Dickie and Rattan [2010]) rejecting Fregean sense *sensu stricto*, but allowing that shared social spaces can be thought of as shared senses, at least insofar as they play a similar theoretical role in terms of “cognitive significance,” of both the existential and substitutional kind.

framework and grid is shared (what is square 1 for you is square 1 for me, and so on), then we can trade on the identity of *that and *that by keeping track of its location within the grid. Given that objects, by their nature, cannot share a spatio-temporal location with another object, then they can be uniquely identified this way (see Evans, 1982). There is no need, the thought goes, for anything more than that. Here is Seemann on this point, which he calls the *Communicative Principle for Joint Contexts* (or CP1):

In order for a speaker to be in a position to inform a hearer of a perceptual fact about a jointly perceived object by means of a demonstrative utterance, speaker and hearer must perceptually know in common the bounded location of the object the speaker is intending to refer to and communicate about. (p. 61)

I hope this now gives the reader a sense of what Seemann is arguing for. To summarise: the claim seems to be that PCK occurs within communicative situations where there is a demonstrative identification of an object. Such communicative situations are, at least when successful, necessarily, by their nature, situations in which the subjects enjoy PCK of the demonstratively identified object. This allows for a type of shared reasoning that would be otherwise unavailable without the use of a conceptualised identity statement: for the agents can now share information about the commonly perceived object without needing to establish that the object they are each discussing is the same — they can simply see that it is. What allows them to see that it is the same object is that the object has a place within a shared spatial framework that uniquely identifies the object for any participant who is using that framework.

There is much that is attractive about this proposal. As we shall see, it lends itself well to an underlying externalist, embodied, and enactivist metaphysics of human persons. And this goes some way to satisfying the transparent-yet-communal, rich-yet-lean desiderata for PCK that we discussed in the previous section. We shall look at this metaphysical underpinning shortly, in the next section. Firstly, though, I want to again raise a note of caution about the approach and its capacity to satisfy the strong epistemological demands put upon it.

Seemann's proposal is that a successfully executed informative intention about a demonstratively identified object provides for the kind of overt communicative act that, as we discussed above, the Gricean was looking for. It is overt, on Seemann's account, because both participants are aware of *which* object is being demonstrated in the sentence <*that is a buoy> because the indicated object has a unique place within a shared space. But, for the Gricean, this can only be half of the story.⁶ For what makes an utterance communicative — and not just a matter of the transfer of information from one thing to another — is that the

⁶ As Seemann himself acknowledges (see p. 97).

hearer recognise the *intention to communicate*, and, further, that the uptake (i.e., acceptance of) the information is reliant on the fact that the hearer recognise the communicative intention. This is to say, for the Gricean, overtness of *communicative intention* is at least as important as overtness of *informative intention*⁷ and it is not clear that Seemann's proposal provides the former.

Imagine, for instance, two strangers standing side-by-side in front of the Turner painting. Neither acknowledges the other. Let's suppose that both think (for whatever reason) that the other is deaf. Person A steps forward to the canvas hovers her finger over the orange mark and says (to herself) <*that is a representation of a buoy>, thinking, incorrectly, that the other can't hear her. Then Person B steps forward and hovers his finger over the orange marks and says (to himself) <*that is a last-minute addition to the painting>, also thinking, incorrectly, that the other can't hear him. There is information about a demonstratively identified object passed here between the two within a shared spatial environment. It seems to fit Seemann's analysis. Moreover, it also pulls the trick of allowing each participant to make the necessary inference to <*that is both a representation of a buoy and a last-minute addition to the painting>. But it doesn't fit the Gricean analysis. It is missing a crucial ingredient: what is missing is the *commonality*. It is the commonality that allows the agents to *commonly know* (and not just mutually know) the identified facts, and also to commonly reason in concert. Appealing to the world to offload the cognitive burden of knowing that one is commonly sharing a perspective or idea with another will only work if the world which is being appealed to somehow contains within it that commonly shared perspective. And triadic geometrical relations are not yet shared perspectives.

Social Space

Seemann is clearly alive to the pressures outlined above, since the third section of the book looks into the underlying "spatial framework ... that orders things relative to the positions of a plurality of perceivers and agents" and that ultimately provides for "a reflective conception of space that allows me to treat a variety of positions in it as occupied by perceivers" (p. 88). The discussion is again very far-ranging, marshalling a ton of empirical evidence from cognitive psychology, social psychology, and social neuroscience, as well as engaging in a synoptic overview of different philosophical and psychological categorisations of spatial thinking and behavioural and motor control. Nevertheless, there are two key ideas that run through this section.

First, there is the full-blown picture that is to emerge from the discussion. This picture is one in which there are rational persons, each with their own standpoint-dependent

⁷See Sperber and Wilson (1996) for a clear discussion of the Gricean programme, as well as the distinction between informative and communicative intentions.

perspectives on a shared space, and where those standpoint-dependent perspectives, held by distinct persons, can sometimes align — specifically in situations of PCK and demonstrative-based communication — such that, when they do align, the agents enjoy factive states of knowledge that they share perceptual attention to an object.

Second, there is the underlying framework that is to support the full-blown picture. This framework is intended to be such that it has the resources to support the provocative claim that each agent in a situation of PCK enjoys a *factive* relation to the other person and to their shared environment, such that they know that the state they are in could not be false (i.e., if one is in that state, then one knows — and knows that one cannot be incorrect about — the fact that the other is in the same state with, *mutatis mutandis*, the same content).

How are we to think of the underlying framework? On the one hand, we might just think of it as a set of geometric spatial relations between three physical bodies: your body, my body and the body of the object that is the focus of our PCK. But this cannot be right. When one jointly attends with another person then the other is a locus of mentality of some sort. They are distinguished as *co-attender* and not as an *object of attention*. If the shared spatial framework is to underpin the factive claim about PCK that Seemann wants to establish — such that each person merely needs to look at the situation they are in to noninferentially understand that it is shared with another — then that framework must contain within it the facts that those factive states are about. Clearly, one cannot be in a factive state unless the relevant facts themselves obtain. And the facts that must obtain in this case are that the agents have a shared attention to the object.

On the other hand, the shared space might be defined by agents who have a *perspective* on the world. But what kind of perspective are we talking about? For, again, it would seem not to be enough if we supposed that the perspective the agents had was one that merely recognised other *bodies*. The framework is the underpinning for *common* knowledge, and common knowledge requires that each agent is able to think of the other as a locus of perspective on the world, and to recognise this about each other. And at this point one might feel the original problem has just re-emerged at the level of the underlying framework.⁸

In a move that is not really flagged until very late in the book, Seemann draws on an enactivist understanding of social space to forestall this type of problem. In

⁸ We can see here something of the dilemma that I mentioned above. The shared spatial framework that enables and supports PCK must either (a) contain within it the rich epistemological facts that sustain the idea that the agents can simply open their eyes and factively see that they enjoy full-blown PCK (in which case, we are no further forward in our investigation, since the problem has just been off-loaded to the framework), or (b) contain a mere set of spatial relations such that when the agents open their ideas, they need to bring with them a range of presuppositions that allow them to see those spatial relations as being imbued with overlapping mental perspectives (in which case, we are again no further forward in our investigation, since the problem has just been off-loaded to those presuppositions).

place of the geometric and extensionally defined spatial relations of “joint perceptual constellations” we get a more dynamic understanding of shared space that is primarily defined, not by external relations in metaphysical “Newtonian” space, but by the possibilities for joint action that are afforded to two embodied entities. Seemann calls this a “social action space...[a] kind of spatial framework [that] is practical, not theoretical; it enables action rather than spatial thinking” (p. 150).

The precise account that Seemann provides of the shared action space is complex and encompasses a discussion of Gibsonian affordances, Gallagherian body schemas, mirror neurons and attentional binding among much else. My understanding of the central claim, however, is that there is a basic form of interaction — exhibited in low-level cues and resonances that emerge at a preconscious level between individuals within a shared space — which allows the cognitive systems of each individual to bind together sub-personal information from early vision and proprioception to provide a sense of location (p. 148). What is not quite clear is how this binding is meant to relate to the other person. At times one feels that Seemann *should* be arguing that (for each individual) low-level visual processing of the other agent’s actions are bound together with one’s own to create an undifferentiated primordial sense of “we” that structures the behavioural space of the participants. This, it seems to me, if developed in detail, might allow for the kind of embodied and spatial sense of intersubjectivity that could then be unravelled and relied upon within PCK such that each agent is, at the personal level, simply in a position to see that they are sharing their attention.

As far as I can tell, however, Seemann hints at, but never quite makes, this argument. Nevertheless, he is clear that social space is defined in terms of a fairly tight “bodily communication,” and he goes on to endorse Colwyn Trevarthen’s primary intersubjectivity account of early social interactions — which take something like this form — so we are justified, I think, in taking the above view as in the spirit (if not quite the letter) of Seemann’s argument.⁹

This now completes the overview of Seemann’s argument. It is a picture of PCK in which PCK is possible within communicative interactions with demonstrative elements that are located within a shared space. This shared space is defined and shaped by low-level “cross-creature embodiment” interactions that enable the participants to understand each other as loci of perception and action within a world that is public and shared.

Joint Attention and PCK

The final, short section of the book pulls together the discussion outlined so far and then relates it to current thinking and evidence in developmental psychology,

⁹In an earlier work, Seemann developed the concept of “cross-creature embodiment” that comes close to the primordial, undifferentiated sense of “we” that I have suggested above (Seemann, 2011b).

especially with regards to children's developmental trajectory. Seemann summarises his view as follows:

... there are two distinct levels to the analysis of perceptual common knowledge. There is, first, the joint interaction in which a practical, embodied relation between agents and environment is established. And there are, second, the agents' mental states of knowing that are individuated by that relation. (p. 179)

There is much of interest that follows in the discussion of joint attention, including an application of the framework to the so-called "developmental paradox" (pp. 189–193).¹⁰ However, my concern for the remainder of this review is to assess how well Seemann's proposal handles the problem that we started with: the relation between joint attention and PCK.¹¹

For all its virtues, I think there are a pair of difficulties that arise for the kind of two-tier, bottom-up approach that Seemann is recommending. The first goes from bottom to top. If we suppose that there is a minimal behavioural core to joint attention from which a fully conceptualised and propositional PCK can develop, then we are liable to run into problems when trying to scale up. How, for example, do infants come to develop a conceptualised understanding of minimal embodied engagements? It is not clear. It can't be via language, presumably, because, on Seemann's picture, basic linguistic communication occurs at that higher, propositional level. The key move that allows for scaling-up must occur elsewhere than in the developmental base.

The second problem goes from top to bottom. If we suppose that PCK is based on awareness of a minimal triadic base — and, moreover, a minimal base that can occur in the absence of the conceptualised PCK — then our initial problem seems to just re-arise, for our capacity to communicate, or our capacity to rationally launch a coordinated action, depends on our *both* being conceptually related to that spatial base. But how am I to establish that? The base is, by definition, free of that information. So, we seem to be back on the iterative carousel.

With both of these problems, the really heavy-lifting has to be going on at a conceptual level. For this reason I think that any theory which attempts to reductively ground conceptual, semantic, and epistemological features in a minimal base that is free of those elements is going to have its work cut out. A more promising approach, in my view, would be to try to explain how infants are born and socialised within a rich conceptual social world that already contains within it,

¹⁰ See de Bruin and Newen (2014) for a useful discussion of the developmental paradox.

¹¹ In fact, Seemann goes on to deny that there is any such thing, strictly speaking, as joint attention, at least if joint attention is considered as a unifying phenomenon that it is often taken to be (p. 164). Instead, he says, there is the lower-level joint perception and higher-level PCK. This, however, is a terminological dispute, and my question can just be reframed, if one wants, as the relation between joint perception and PCK.

almost from the beginning, the pedagogic scaffolding for genuine communicative actions and full-blown perceptual common knowledge.¹²

PCK is janus-faced. It faces towards the world and it faces towards other people with whom one shares the world. Any account of PCK or its cognates needs to find a balance between these two aspects. The guiding idea behind Seemann's approach is that a lot of social understanding is achieved just by dint of the fact that we are gazing out, and acting within, an external world that is mutually accessible to oneself and to others. There is a lot of mileage to be had with this approach, as the richness of Seemann's text attests. But it also has its limitations. For it is not just the world that needs to be mutually accessible, it is also each other: our intentions, our perspectives, and our background assumptions. One might wonder, then, whether the minimal enactivist base that Seemann posits has the resources to lift mutual knowledge into common knowledge, or an informative intention into a communicative intention.

The Shared World is a rich and energetic book, that deserves a wide readership. There is so much going on in it that it is sometimes difficult to separate the main lines from the side alleys. Yet, for all that, it is a very impressive work of critical synthesis, that manages to construct an empirically informed, yet philosophically astute, picture of the basis of human sociality. Even for those, like myself, who might take issue with some aspects of the argument, there is, within these pages, a lot to learn from, much to agree with, and even more to engage with. It will be interesting to see where Seemann intends to take this framework next.

References

- Davidson, D. (2001). The second person. In D. Davidson, *Subjective, intersubjective, objective* (pp. 107–121). Oxford: Oxford University Press.
- de Bruin, L., and Newen, A. (2014). The developmental paradox of false belief understanding: A dual-system solution. *Synthese*, 191, 297–320.
- Dickie, I., and Rattan, G. (2010). Sense, communication and rational engagement. *Dialectica*, 64, 131–151.
- Evans, G. (1982). *Varieties of reference*. Oxford: Oxford University Press.
- Gilbert, M. (1989). *On social facts*. Princeton: Princeton University Press.
- McCulloch, G. (1984). *The game of the name: Introducing logic, language and mind*. Oxford: Oxford University Press.
- Schiffer, S. (1972). *Meaning*. Oxford: Oxford University Press.
- Seemann, A. (Ed.). (2011a). *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience*. Cambridge, Massachusetts: MIT Press.
- Seemann, A. (2011b). Joint motor action and cross-creature embodiment. *Review of Philosophy and Psychology*, 2, 279–301.
- Seemann, A. (2019). *The shared world: Perceptual common knowledge, demonstrative communication, and social space*. Cambridge, Massachusetts: MIT Press.
- Sperber, D., and Wilson, D. (1996). *Relevance: Communication and cognition* (second edition). Oxford: Blackwell.

¹²I take this kind of approach in Wilby (2020).

- Syverson, P. (2003). *Logic, convention, and common knowledge*. Chicago: CSLI Publications.
- Wilby, M. (2020). Modest sociality, minimal cooperation and natural intersubjectivity. In A. Fiebich (Ed.), *Minimal cooperation and shared agency* (pp. 127–147). Cham, Switzerland: Springer Publishing.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.