

Pointing and Representing – Three Options

Nick Young[†]

Nick.Young@ua.ac.be

Angelica Kaufmann[†]

Angelica.Kaufmann@ua.ac.be

Bence Nanay^{†‡}

Bence.Nanay@ua.ac.be

ABSTRACT

The aim of this paper is to explore the minimal representational requirements for pointing. One year old children are capable of pointing – what does this tell us about their representational capacities? We analyse three options:

- (a) Pointing presupposes non-perceptual representations
- (b) Pointing does not presuppose any representation at all
- (c) Pointing presupposes perceptual representations

Rather than fully endorsing any of these three options, the aim of the paper is to explore the advantages and disadvantages of each.

KEYWORDS: Pointing; Nonperceptual representation; Perceptual representation; Antirepresentationalism; Attention; Epistemic states.

1. Introduction: Pointing and Representing

1.1 What is pointing?

Human beings share a species-universal ability for gestural communication. The functioning of pointing gesture in humans is a widely investigated field.

[†] Centre for Philosophical Psychology, Philosophy Department, University of Antwerp, Antwerp, Belgium.

[‡] Peterhouse, Cambridge University, UK.

Yet, the ontogenetic roots of this behaviour remains largely unknown. Vygotsky (1978) hypothesized that pointing develops from failed attempts at reaching towards objects. More recently, however, it has been noted that infants continue trying to reach for things well after mastery of pointing has been achieved (Masataka, 2003), suggesting that these abilities could have developed independently. According to another proposal (Bates et al., 1975), pointing originally emerges for non-communicative purposes and only later in ontogeny does it become a communicative device. However, other researchers have shown that pointing performance is significantly more frequent when it occurs in social contexts (Franco et al., 1996). A third proposal (Bruner, 1983) is that the development of pointing is triggered by social practice. These lines of analysis are not mutually incompatible. Still, there seems to be no agreement on which account, if any, is capable of providing a full explanation of the ontogeny of pointing.

An aspect of the research on which there seems to be more agreement is the age at which infants master pointing. Recent findings report that, from the age of 12 months, infants begin to show a bidirectional appreciation of the role of pointing gesture (Behne et al., 2011). This appreciation becomes fully evident between 14 and 18 months of age. How young infants are sensitive to others' pointing gestures relates to their own pointing gesture performance: infant pointing comprehension and production mutually correlates (Liszkowski et al., 2006; Liszkowski, Carpenter & Tomasello, 2008; Tomasello, Carpenter & Liszkowski, 2007). In this paper we will not analyze these correlations. Our focus is the production, not the understanding, of pointing.

Pointing is a device capable of directing the attention of others. Even though pointing can allow for reference displacement (making it a good candidate for a precursor for linguistic communication) it is a communicative mode that fully depends on sharing the same visual space (Liszkowski, 2010). Through pointing, infants can direct the attention of others towards an object of their choice. Thus, it seems that infants take some other beings to be potential attenders and they appreciate that their attention can be influenced. The question is: what kind of representations (if any) are needed in order to account for this?

We examine three possible accounts of the representational mechanisms responsible for pointing production:

- (a) Pointing production requires the non-perceptual representation of the other person's mental states (of, for example, whether she is attending).

- (b) Pointing does not presuppose any representation at all.
- (c) Pointing does presuppose some representations, but only perceptual ones.

Before we consider the advantages and disadvantages of these three options, we need to outline two desiderata that any theory of pointing needs to be able to satisfy.

1.2 Two desiderata for any account of pointing

Infants have been found to point for various purposes including: to obtain a desired object or to achieve a desired effect, to share interest with another person, to provide another with helpful information, or to request information about the world. Here, we shall focus on two features common to all of these types. Any account of pointing behaviour must be able to satisfy the following desiderata:

- (1) Any account of pointing must be able to explain that the person who is pointing responds to the other person's line of sight.
- (2) Any account of pointing must be able to explain that the person who is pointing appreciates the epistemic states of the other person.

(1) should not be taken to mean that the infant explicitly understands or represents what the other person is perceiving – this would automatically rule out any anti-representational account of pointing. Rather, this 'responding to the other's line of sight' can be fleshed out as sensitivity to the orientation of the other with respect to the environment. That is, pointing production is sensitive to whether the other's eyes are oriented in the direction of the object to which the pointing gesture refers. But this is not all that is meant by this desideratum. Responding to the other person's line of sight also means that if a point is unsuccessfully followed and the other is oriented towards the wrong object, the infant will point again. The infant is also able, to a degree, to control this spatial relation. The direction in which the infant's hand moves will cause the other to move their head, eyes, or body so that the trajectory of their line of sight is changed. In short, pointing is (i) sensitive to the other's line of sight prior to pointing production, (ii) sensitive to the other's line of sight during and after pointing production and (iii) used to change the other's line of sight.

Pointing, then, demonstrates a remarkably fine-grained responsiveness to another's line of sight – much finer grained than other forms of non-verbal communication. Vervet monkeys, for example, can act so that conspecifics look in one direction or another but do not direct their gaze to a particular spatial location: on hearing a 'snake alarm call', they look down; on hearing an 'eagle alarm', they look up; and on hearing a 'leopard alarm'; they run up into the trees (Cheney & Seyfarth, 1980). In this case, although different types of alarm call will cause other monkeys to look to the sky, the ground etc. they lack precision. There is no call for 'snake directly below the tree I am in' or 'eagle at five o'clock'. The trajectory of an infant's point on the other hand, controls precisely where the other should direct their gaze.

The second feature of pointing is the appreciation of the epistemic states of others. Liszkowski et al. (2008) conducted a study in which 12 month olds watch an experimenter place an object in a specific location, shortly followed by that object being moved to a second location. In some cases the infant saw the experimenter watch the object being moved, in others they saw that the object moved to the second location when the experimenter was looking elsewhere. When the experimenter did not see the object being moved the infants were more likely to point to the object in its new location than they were if the experimenter watched the displacement take place. Such a result suggests that pointing infants are able to keep track of what others know and do not know.

The following three sections will assess how well three different types of cognitive framework can account for these two aspects of pointing. In Section 2, we analyse the account according to which pointing production requires the non-perceptual representation of the other person's mental states. In Section 3, we examine the proposal that pointing production does not require any representation at all. And in Section 4, we assess the plausibility of an account that takes the middle way between these two approaches, according to which pointing production does require representations, but only perceptual ones. The aim is not to give a knock-down argument in favour of one of these three approaches and against the other two, but rather to consider the advantages and disadvantages of each.

2. First option: Non-Perceptual Representation (NPR)

What does it mean to represent something? What does it mean to have a mental representation? A very common way of using the concept of mental representations is this: representing means attributing properties to entities. If I represent the carpet as red, I attribute a property ('being red') to an entity (the carpet). Mental representations are thought to play an important role in explaining actions. For example, a person's careful handling of a vase might be explained by their attributing to it the property 'valuable' or 'my grandmother's favourite vase'. Another way to put this would be to say that they represent the vase as being valuable or as being my grandmother's favourite vase.

Some representations are perceptual representations, whereas some others are non-perceptual representations. Perceptual representations attribute properties perceptually; non-perceptual representations attribute properties non-perceptually. Another way of putting this would be to say that some, but not all, properties we attribute to entities are perceptually represented. The colour or the shape of the vase are perceptually represented properties. Being my grandmother's vase is not a perceptually represented property: we cannot see the vase as being my grandmother's favourite vase.

Many philosophers explain actions in terms of non-perceptual representations: I believe that it is raining outside, I want to avoid getting wet, I believe that the best way of achieving that is to take an umbrella; therefore, I form an intention to take an umbrella. All of these mental states are non-perceptual representations. Explaining action in this manner is widely used by philosophers, cognitive scientists and everyday folk alike (Sterelny, 1990; Davidson 2001). Importantly, it is also a common type of explanation in social cognition, with many theories seeking to explain our understanding of others, through the attribution of mental states to the other (for example, Carruthers & Smith, 1996, Nichols & Stich, 2003, Goldman, 2009). The attribution of mental states to others is, at least, according to the mainstream conception, a non-perceptual attribution of properties such as 'having a desire', 'having the intention', 'having a belief' or 'having some other types of mental state'. As Goldman summarizes, "to attribute a mental state to an individual is to represent that individual as being in that state" (Goldman 2009, p. 235). This section will assess how well this type of cognitive framework (which we call

NPR for Non-Perceptual Representation) can be used to account for various features of pointing production.

2.1 NPR and the content of representations required for pointing

The first thing that an NPR account of pointing must do is specify precisely what property the infant is attributing to the other. When doing this care must be taken not to overstate the cognitive capacities of 12 month olds. Two approaches to social cognition which do involve the attribution of unobservable mental states are Simulation Theory (Gallese & Goldman, 1996, Goldman, 2006) and Theory Theory (Nichols & Stich, 2003), both of which have been put forward as theories which can account for the sophisticated social skills which adults demonstrate. To take an example from Dennett, if we can watch a film and, effortlessly, think of one character “He wants her to think he doesn’t know she intends to defraud her brother!” (Dennett, 1989, p. 48) then such a complex attribution requires quite sophisticated representational powers. If however, similar theories are to be used to explain the emergence of pointing in infants, then the limited cognitive abilities of such young children must be taken into account. Pointing normally emerges at around 12 months and infants have developed some social cognitive abilities, including: an appreciation of attention to the self (Reddy 2005); recognition of goal-directed action (Csibra 2008) and some appreciation of others’ false beliefs (Kovacs et al., 2010). However, 12 month-olds cannot yet speak in sentences (see Butterworth & Morissette 1996) and fail to appreciate others’ false beliefs, at least in some situations (Wellman et.al. 2003).

In light of this, there is broad agreement among those who advocate NPR theories of social cognition that the representations involved in joint attention behaviours must be of a more minimal type than those invoked by adults. There is also some agreement as to the type of representational state which meets this requirement. Joint attention behaviours are often held to involve an infant’s representation of the *perceptual experience* of the other (Hutto 2008, p. 127 – 128; see also Nichols & Stich, 2003, p. 135 – 136.). As pointing involves the other’s attention being directed, and thereby their perceptions being altered, representing perception seems a plausible way to account for this behaviour. It is also thought to be suitably minimal, as representation of others’

perceptions is considered less cognitively demanding than representing their beliefs, or their epistemological states.

2.2 Representing Perception and Responsiveness to Line of Sight

Representing the other's perceptual state would seem to provide a good explanation for the infant's appreciation of line of sight relations between other and object. As we saw in Section 1, pointing behaviours indicate that infants can keep track of whether the other is correctly oriented towards the target object. If the infant is representing the perceptual experience of the other we can see how this is possible. They are able to represent what the other perceives before the pointing gesture is made (i.e., not seeing the target object), represent what they want the other person's perception to be (seeing the target object) and then point as a means of making this come about. If the other turns towards the 'wrong' object, or does not turn at all, the infant represents the other's perceptual state, matches it against the perceptual state which the infant desires the other to have, and seeks to rectify the situation by pointing again.

However, when it comes to the infant's appreciation that the other's line of sight can be controlled by pointing, it is more difficult to see how an explanation is possible solely in terms of representations of perception. To see the problem, we need to focus on the infant's situation before it has made the pointing gesture. If it is correct that the infant is able to represent perceptual experience it will represent the other's current perceptual state as *not* perceiving the target object. It will also be able to represent the perceptual experience which it wants to elicit in the other, that is, perceiving the target object. It can also represent another possible perceptual state of the other: perceiving the infant's point. However, simply being able to represent this last state is not sufficient for the infant's understanding that point in a particular direction will *cause* the other to turn their head. Neither does it seem sufficient to explain how a point in a particular direction will lead to a particular turn of the head. Representations of perceptions give no information about how a point would be interpreted by the other and so to understand others' responsiveness to points would seem to require more than just representation of perception on the part of the infant. There are ways of addressing these worries that we will explore in the next subsection.

2.3 Representing perception and appreciation of knowledge and ignorance

Neither does representing the perceptual states of others seem sufficient for explaining how infants can appreciate another's knowledge or ignorance of an object's location. In the Liszkowski et al.(2008) experiment, the difference between the knowledge and ignorance conditions is not what the other person (the experimenter) perceives now, but what she perceived *before*. And representing what another agent perceived before is not representing a perceptual state but representing a non-perceptual state. In the face of this difficulty, and the one concerning the first desideratum we outlined in the previous subsection, a defender of NPR could follow one of the following three approaches:

- (i) Accept that mental states other than perceptual ones can be represented
- (ii) Argue that besides representing the other person's perceptual state, the representation of some non-mental properties is also needed
- (iii) Extend what can be contained in perceptual content.

(i) Proponents of NPR could simply bite the bullet and allow that pointing depends on representing the other's perceptual state *and* representing some other mental state of theirs as well. As we have seen, representation of perception is preferred by many NPR theorists as it is considered to be not particularly mentally demanding and so minimises the risk of over-intellectualising the infant's cognitive powers. However, there is some reason to think that mental states other than perceptual ones are represented around the time pointing emerges. In the last few years, some have argued that infant's understanding of 'higher' mental states in others such as belief emerge much earlier than previously thought. Kovacs et al. (2010), for example, suggest that a registering of others' beliefs emerges as young as seven months. Furthermore, various studies show that infants younger than 12 months have some appreciation of goal directed action (Gergely & Csibra, 2003; Csibra, 2008). If it is the case that infants can represent beliefs about the world then this would allow for a clear NPR explanation of why pointing occurs more often when the other person is ignorant of the target object's location: the infant represents them having an incorrect belief about where the object should be, and points so as to correct this. It is less obvious, however, how an ability to represent beliefs could lead to the infant's ability to direct another person's line

of sight. However, if one year olds are able to attribute more sophisticated mental states to others, this can only help NPR based explanations of pointing

(ii) A second possible move would be to admit that pointing involves more than just representing the other's potential and actual perceptual states. However, rather than allowing for extra mental states to be attributed to the other, the behaviour could be explained in terms of the infant attributing a perceptual state and a *non-mental* property to the other. Although an infant's cognitive capacities at 12 months are limited, they do exhibit some understanding of inanimate, physical objects in the world. They, for example, have some appreciation of gravity and are surprised when round objects are seen to roll up rather than down an inclined plain (Kim and Spelke, 1999). An NPR explanation of this type of understanding could be made in terms of attribution of basic physical properties to these objects such as 'rolls downhill when unsupported'. Although less useful in explaining infants' ability to track the knowledge or ignorance of others, a proponent of NPR could argue that the infant's controlling the other's line of sight is dependent on it attributing to the other a physical property of this type. In the same way that an infant's deliberate dropping of a cup is driven by its attributing 'falls to ground when unsupported', its pointing so as to cause a head turn is driven by its attributing 'turns in the direction of a point' to the other person. On this explanation, the infant would still represent the current and desired perceptual state of the other, but would rely on the attribution of a physical property to effect the change from one to the other.

(iii) The third, and perhaps most ambitious option for NPR explanations of pointing would be to extend what counts as perceptual content, and thereby extend what can be included in the infant's representation of the other's perceptual state. As we shall see in Section 4, some philosophers hold that as well as obviously perceptual properties, such as shape or colour, perceptually attributed properties might also include: causal efficaciousness, presenting action opportunities, dispositional properties and normative or prescriptive value. Siegel, 2005, 2009; Nanay, 2011a, 2011b 2012a; Kelly, 2010; Matthen, 2010) The dominant (Hutto, 2008, p. 127 – 128), although not the only (Nichols & Stich, 2003, p. 135 – 136), view on representing the perceptual states of others, is that the content of this type of representation is partly derived from the infant's own perceptual states. If this is accepted, and the infant's own perceptual states involve one or more of these non-obvious perceptual properties, then the perceptual states it attributes to others might

include them also. Allowing this would expand the options available for the proponent of the NPR picture and so might allow for pointing to be explained solely in terms of the attribution of perceptual states. The disadvantage of pursuing this approach however, is that claiming causal efficaciousness, prescriptiveness etc. as perceptual properties is not uncontroversial. Furthermore, simply allowing these properties into the infant's representations of others' perceptual states is only half the battle. In similar fashion to (i), more work would be required to show how representing these extra perceptual properties would allow for the infant to control the other's line of sight.

3. Second Option: Anti-representationalism (AR)

As we have seen, according to NPR, pointing can be explained with reference to non-perceptual attribution of properties to entities. However, in the last twenty years or so, a competing cluster of theories have emerged which argue that intelligent interaction with the environment need not be driven by mental representations at all. We will here introduce two ways in which Anti-Representational theories (AR) can explain human behaviour and evaluate how well they can be used to explain pointing production.

3.1 Anti-representational approaches and social cognition

In lieu of representations, a number of AR theories endorse complex, reciprocal cycles of causation between person and environment as a means of explaining human behaviour (Thompson, 2007). Whereas NPR theories consider action and perception to be discreet categories (an object is seen, its unobservable properties determined, then acted upon) AR theories hold that the interplay between our perceiving the world and acting upon it “can be so subtle and complex as to *defy* description in representational terms” (van Gelder, 1997). If I pick up grandmother's vase and try to find the most secure way of holding it, the movements of my fingers make will be guided by the tactile sensations caused by the vase's weight and shape. The type of sensory feedback received will give rise to further movement, which will in turn give rise to further sensation. Here, the weight and shape of the object are not represented, but instead help to determine the particular patterns of dynamic interaction (see Thompson 2007, p. 13-14).

A similar account can be seen in social cognitive variants of AR, which, drawing from mirror neuron studies, argue that our ability to skilfully engage with others derives from a “primordial”, empathetic connection with other people. This link is considered the basis for very young infants’ ability to take part in dyadic interactions between themselves and another person. Before they can point infants will mimic facial expressions of others, attempt to make eye contact, or –depending on mood– hide from the other’s gaze in an apparent display of coyness (Meltzoff & Moore, 1977; Reddy, 2000, 2003; Trevarthen & Aitken, 2001). The fluidity and reciprocal nature of these ‘protoconversations’, along with avowals that such abilities are representation-less (Zlatev et al., 2008, p. 23), suggests an approach to social cognition very similar to the type of dynamic interaction endorsed by AR.

A second feature common to AR theories is the idea that behaviour in a given situation is influenced by all aspects of our immediate environment and should not be thought of in terms of a simple one-to-one relation between person and object (Seemann, 2010; Thompson, 2007; Dreyfus, 2002). If I am walking along a bumpy road, the argument goes, I do not need to attribute properties to every single bump in order to navigate them. Instead, the presence of these features in my environment will influence my behaviour generally, such as my gait or speed of walking. The AR theory would argue that our behaviour is driven by this type of what we might call ‘holistic sensitivity’ to our environment, rather than representation. Suppose for the sake of argument that we can make sense of this concept of holistic sensitivity and do so that would not make this holistic sensitivity a variant of unconscious representation. The question then is how this AR account could be applied to social cognition and, more specifically, to pointing production.

If this idea is applied in social cognition we can think of other people as a particular type of environmental feature. Just as the vase can influence behaviour without us representing it, the very presence of another person in my immediate environment can affect the type and manner of interactions made. Furthermore, if we are primordially connected with one another then we might expect other people to constitute especially salient features of the environment, the proximity of which can influence how we behave in particular and unique ways.

3.2 AR and the two desiderata

The notion of holistic sensitivity can be used to provide an AR account of the pointing infant's appreciation of line of sight. The idea that we have a primordial connection with others has been developed by philosophers such as Seemann (2010, 2012), Gallagher (2012) and Campbell (2002, Chapter 8) so that it is not simply the presence of another person in a shared environment which can influence interaction, but also the direction of their gaze. Seemann (2010) gives the example of a driver with a passenger sitting next to him in which the direction the passenger looks can affect where the driver is looking, as well as his behaviour more generally; the passenger glancing to the left might compel driver to look left also, the passenger peering to look into an adjacent lane could cause the driver to change lanes. An AR account of this situation would argue that the driver's appreciation of his passenger's line of sight does not involve the attribution of perceptual states, mental states of any other type of property. Instead, line of sight appreciation is cashed out in behavioural terms: the driver's sensitivity to one aspect of his environment (the direction of his passenger's eyes) is factored into how he behaves. An infant's ability to appreciate whether the other person is directed towards the target can be explained in the same manner: it does not represent the other as perceiving anything, but instead is sensitive to whether or not their eyes are aligned appropriately with the target. Once this line of sight relation has been registered the infant's subsequent behaviours will be modified.

One worry with this account is that the same story can be told in a representationalist framework: if the passenger glances to the left, the driver unconsciously represents this head movement and it is this unconscious representation that is responsible for why the driver also looks to the left. The considerations for the AR account of social cognition are considerations mainly fuelled by the insight that we are often not aware of features of our social environment that nonetheless influence our social behaviour. But these considerations fail to rule out that we do represent these features of our social environment but do so unconsciously. Representation may or may not be conscious and it seems that AR accounts jump from the lack of conscious representation to the lack of representation *per se*.

The notion of dynamic interaction can be used to provide a AR account of how infants control line of sight relations through pointing. Although dynamic interaction emphasises reciprocal cycles of causation between person and

environment, supporters of AR theories are keen to stress that action should not be understood as mere passive response to whichever features of the environment are in the vicinity. Instead, we might think of the agent as being able to initiate different patterns of reciprocal causation with the same environmental feature depending on their present situation. A vase can be picked up, pushed over or thrown depending on the situation and we can think of an infant having a similar repertoire of skills which she can utilise in the presence of another person. In some cases it might initiate reciprocal dyadic interactions, such as imitation or holding eye contact, in others, such as when there is a desirable or interesting object around, it will respond to the other by pointing. If an infant's appreciation of the other's line of sight can be understood in terms of how its behaviour is influenced by an 'other directed towards object' feature of the environment, its ability to control the other's line of sight through pointing can be thought of as its knowing how to bring forth such a feature. The tight connection between environmental features and a person's responses included dynamic interaction could be used to explain the ease with which an infant directs the other's gaze to a precise location in their environment. In the same way that the contours of the vase guide my fingers as I grasp it, the particular spatial configuration of person and object can guide the angle and direction of the pointing gesture.

Thinking of pointing as a deployable skill goes some way to explaining how infants recognise others as potential attenders. We saw in the previous section that accounting for pointing behaviours purely in terms of the attribution of perceptual states to the other leads to difficulties for NPR, as it is not clear how the infant would know that a particular movement would lead to a particular response from the other person. We also saw that one way this difficulty might be avoided would be to argue that in addition to perceptual states, infants rely on their knowledge of how they can manipulate physical inanimate objects. The AR account of pointing outlined here is similar in that it also relies on the infant knowing how to effect physical changes to its environment. Where the two accounts differ is firstly that AR holds the infant's ability to control line of sight consists entirely of its practical understanding of how to effect changes in the world without representing the other's mental life. Secondly, as we have seen, the infant's ability to interact with the world is not based on the attribution of properties but instead by initiating cycles of dynamic interaction with objects. Therefore, AR would account for the infant's recognition of others as potential attenders by saying that the other person is recognised as

the type of environmental feature which can be manipulated through a particular type of dynamic interaction.

The problem here is that it is unclear whether talk of ‘dynamic interactions’ addresses the main worry that faces an AR explanation of pointing production in general and of the problem of how infants recognize potential attenders in particular. Presumably, these ‘dynamic interactions’ will be very different if the other entity is the infant’s caregiver or an inanimate object. But then what explains this difference in a way that would not appeal to the concept of representation? How can the AR account explain this difference? Appealing to ‘dynamic interactions’ is unlikely to work here as the question is precisely why one has different kinds of ‘dynamic interactions’ in the face of these two different entities.

Finally, it is even less clear how AR can be used to explain infants’ sensitivity to what other people have and have not seen. Both dynamic interaction and holistic sensitivity are concerned with the currently perceivable features in a subject’s current environment, but it is not obvious how either could be used to explain the fact that an entity’s history can affect the type of interactions it elicits. I could replace my grandmother’s vase with a duplicate which is perceptually identical but lacks the original’s history and with it, any sentimental value. It is plausible to think that this switch will have an impact on my behaviour; I might be much less concerned about breaking the vase and so will not move around the room with the care that I once did. Similarly with pointing, a person will look the same regardless of whether they have encountered a particular object in the past but an infant will still respond to them differently.

AR theories do provide some account of how different responses can be learned for the same stimulus. Dreyfus suggests that the behaviour elicited by a feature of the environment can be “tuned” (2007, p. 261) by previous successful and unsuccessful interactions. The precise bodily movements which lead to success are reinforced, and those that lead to failure are discarded. It could be argued that replacing grandmother’s vase with a duplicate will in fact *not* lead to an immediate change in behaviour; I might be so used to taking care around a vase of such and such a colour and shape, placed on *that* particular spot on the mantelpiece that I automatically take care around it, despite its lack of worth. Perhaps over time my behaviour will slowly adjust, as carefully walking around a worthless vase is realised to be a non-optimal way of behaving.

This option however, is not available for an AR account of knowledge and ignorance appreciation. In the Liszkowski studies (2008), the infant witnessed the other losing track of where the object was and immediately responded by pointing to the object in its new location. This therefore would not be a case of behaviour being slowly fine-tuned but a wholesale and instantaneous re-tuning. However, this strategy is not a particularly appealing one for the AR proponent as it is quite far removed from the theories main aim of explaining behaviour in terms of subtle sensitivity and response to the environment. Moreover, it is unclear whether instantaneous re-tuning is significantly different from the NPR strategy of property attribution.

4. Third Option: Perceptual Representation

4.1 Perceptual representations and social cognition

A third way in which interaction with the environment can be explained is in terms of Perceptual Representations (PR). Similar to NPR, this framework would posit representations that attribute properties to entities (Burge, 2010, Chapter 2; Nanay, 2010; Crane 2009). NPR, as we have seen, involves the non-perceptual attribution of properties to entities. PR, on the other hand, amounts to the perceptual attribution of properties.

If we assume that we have representations, these representations attribute properties to entities. But this attribution can be perceptual or non-perceptual. Some examples for properties that are attributed perceptually include colour, spatial location or shape. Some examples for properties that are attributed non-perceptually include ‘being my grandmother’s vase’, ‘being made in China’ or, in the case of properties attributed to people, ‘believes that X’.

PR is the proposal that the minimal representational resources needed for pointing are perceptual representations. PR contrasts with NPR in that the representations that make pointing possible are perceptual (rather than non-perceptual) representations. As perceptual representations are generally taken to be phylogenetically and ontogenetically more basic than non-perceptual ones, PR may be able to deal with the over-intellectualization worries better than NPR.

PR also differs from AR inasmuch as it does posit representations in order to explain pointing. As a result, it may be able to account for those desiderata that AR was struggling to cover. The general picture is that PR is the moderate

option between the two extremes of NPR and AR. Interestingly, PR is much less commonly used than AR or NPR in studies of social cognition.

4.2 Representing the other agent perceptually

We will consider two different versions of PR here. The first one is Juan Gómez's account (Gómez, 2005, 2007). Gómez argues that apes and young infants have a "sensorimotor" understanding of whether other people are in "the state of being oriented to potential targets of action". He argues that, unlike belief or knowledge, when a person is attending to an aspect of the environment they display various perceivable features such as characteristic orientations of head, eyes or body as well as an accompanying environmental feature "at the other end" (Gómez, 2005, p. 74). Furthermore, Gómez is quite explicit in saying that the appreciation of these external expressions is a case of representation: the properties we perceive others as having can be integrated into plans for action, and are not simply the product of anti-representational, reflex-like mechanisms (Gómez, 2005, p. 73). The general picture is that we perceptually represent the other agent as attending – because attending is a perceptually salient feature of agents. Gómez then goes on to argue that in only this perceptual representation is needed in order to explain pointing. For pointing, we need to represent the attentional state of the other agent, but we do so perceptually.

Gómez's account could be thought of as a version of the 'direct perception' accounts of social cognition. It has been argued that one of the most basic forms of social cognition involves merely seeing the mental states of another person on her face (Gallagher, 2005, 2008; Zahavi, 2008; Ratcliffe, 2007; de Bruin, Strijbos and Slors, 2011). While some of these authors would prefer to think of this direct perception of the other person's mental state in non-representational terms (leading to a version of AR), if perception is thought of as a representational, as Gómez claims it is, then this proposal would amount to one way of cashing out PR. In short, the suggestion is that there is no need to non-perceptually attribute mental states of others are directly perceivable.

A possible worry about this view comes from philosophy of perception: why should we take 'attending' to be a perceptually attributed property? As we have seen, there is a debate about how to draw the line between perceptually and non-perceptually attributed properties and there is a disagreement about whether sortals or dispositional properties are perceptually represented

(Siegel, 2006; Nanay 2011a, 2011b, 2012a, 2012b, 2013). But the property of attending seems like a much less perceptual property than any of these contested ones. The worry then is the following. Rather than *perceiving* the other as attending, the displayed features of other people are used to *infer* that she is attending.

Even if we accept Gómez's claim that attending is normally accompanied by perceptually salient markers, it does not follow that attending *itself* is perceptually represented. It is still possible that while the accompanying perceptually salient markers (bodily position, orientation of the head, etc) are perceptually represented, and this perceptual representation of these markers leads to the non-perceptual representation of the agent as attending. Because of these worries, Gómez's account (and the 'direct perception' accounts of social cognition in general) would need to be supplemented by an argument that would show that the property of 'attending' (or, 'feeling pain', etc) are perceptually attributed. Without such an argument, these accounts are incomplete.

4.3 Representing the object perceptually

A common assumption between NPR and Gómez's version of PR is that the representations in question attribute properties (perceptually or non-perceptually) to the other agent. The entity that properties are attributed to is the other agent. But this is not an necessary assumption of PR (or NPR). Another way to flesh out these representational accounts would be to say that the representations attribute properties to the object we are pointing to. This is exactly the option explored by the other version of PR we consider here: the theory of Vicarious Perception (Nanay, 2013, Chapter 6).

According to Bence Nanay, one of the most rudimentary forms of social cognition can be explained by the simple perceptual process of Vicarious Perception: of perceiving an object as affording an action to another agent. Note that while the theory of Vicarious Perception talks about perceptual representation, it is not the perceptual representation of the other agent (as attending), but the perceptual representation of an object as affording an action to the other agent.

Here is an example. I am sitting in my armchair looking out of the window. I see my neighbour running to catch her bus. There are lots of people in the street and my neighbour is zigzagging around them on her way to the bus that

is about to leave. How will I represent the lamppost in my neighbour's way? I will not see it as affording an action to me: I am not about to perform any action, let alone an action that would involve the lamppost. But I don't see it in a detached, action-neutral way either: I see it as affording an action (of bumping into) to my neighbour.

The proposal is that we perceive things vicariously all the time. For example, when we are following a sport event on TV, say, a football game, we see the ball as affording various actions for a certain player. We also often perceive vicariously when we are watching movies. Suppose that you are watching a Tom and Jerry cartoon, where Tom is chasing Jerry, who gets stuck in a corner where a hammer happens to be lying around. You see the hammer as affording a certain action not to yourself, but to Jerry (or to Tom, depending on who you're identifying with).

Is the theory of Vicarious Perception better off than the theories of 'direct perception' from a philosophy of perception point of view? Nanay would argue that it is: the property of affording an action to someone else (or the property of being edible for someone else) is the one that is perceptually attributed and we have good reason to accept that the property of affording an action to myself (or the property of being edible for myself) is perceptually attributed (see Nanay, 2011a, 2011b, 2012a, 2012b, 2013). But if so, then it is not implausible to claim that a variant of this property, that is, affording an action to someone else (rather than to myself) and being edible for someone else (rather than for myself) is also perceptually attributed.

But in order to see this, more needs to be said about this property of affording an action to someone else (or being edible for someone else, etc.). The first thing to note is that it is a relational property that cannot be fully characterized without reference to both the object and the other agent's action. Does then the attribution of this relational property to the object presupposes some kind of representation of the other agent? The answer is that sometimes it does, sometimes it doesn't. I can engage with someone else this way without attributing any mental (or even perceptual) state to her. Here is an example: I am still sitting in my armchair looking out of the window. I see my neighbour standing in the middle of the street, deep in thought and I also see that unbeknownst to her, a tiger is approaching her from behind and it is about to attack her. Again, I represent the tiger in an action-oriented manner, but not as affording an action (of running away, presumably) to myself, but to my neighbour. This does not presuppose attributing any properties to my

neighbour. One could argue that I attribute the mental (perceptual?) state to the neighbour that she is *not* aware of the threat and I compare this mental state (attributed to her) with my own mental state of being aware of the threat and this comparison triggers my reaction. But this explanatory scheme presupposes extremely complex mental processes in order to explain a very simple and instinctive reaction.

But Vicarious Perception *can* entail the attribution of properties to another agent. Here is a potential example. Suppose that you are looking at your child carrying a full glass of water very carefully from the kitchen to the living room. You see that the glass is slightly tilted – the water is about to be spilt. In this case, your vicarious perception of the glass may entail the attribution of some properties to your child. But it is important to note that what is attributed here are perceptually salient properties of acting in a certain way – holding the glass this way, etc. Vicarious Perception may also presuppose some kind of representation of agency: that the other person is an agent and not an inanimate object (we don't normally see one object as affording an action to another object). But, as the Scholl and Tremoulet studies show, the attribution of animacy is “fast, automatic, irresistible and highly stimulus-driven” (Scholl & Tremoulet, 2000, p. 299). So if Vicarious Perception does presuppose the representation of agency, this representation is also a perceptual representation.

Now, on to pointing: the claim here is that the minimal representational capacity needed for pointing is Vicarious Perception. When I point at something, I need to perceptually represent the object I am pointing at as affording an action to another agent: I need to perceptually represent it as reachable or out of reach for the other agent. Different kinds of actions are in play in imperative and declarative pointing: when it comes to imperative pointing, it can be any kind of action (i.e. the action I want you to perform with the object), but in the case of declarative pointing, it is normally an epistemic action. Given that we have reason to believe that at least some primates are capable of perceiving vicariously, this may explain why at least some forms of pointing are common among these non-human animals (Leavens et. al., 2005; Leavens, 2012).

4.4 PR and the desiderata

How can PR explain the two desiderata that any theory of pointing needs to be able to account for? The first data point was that pointing behaviour is sensitive to the appreciation of the other person's line of sight. To put it simply, if the other is already looking at the object, even one year olds point to it much less frequently.

How can PR explain this? Gómez's version has a nice way of incorporating this desideratum as the other person's line of sight is one of the perceptually represented features of the other agent (besides the orientation of the head and bodily position). Thus, for Gómez, it is exactly the perceptual representation of the other person's line of sight that makes pointing possible.

Can Vicarious Perception appeal to the same considerations? It can – without endorsing Gómez's controversial claim according to which perceiving the eye of sight (and the orientation of the head, etc) is identical to perceiving the other's attentional state. As we have seen, Vicarious Perception often (but not always) presupposes the perceptual attribution of properties to the other agent. The perceptual attribution of the line of sight can be an important instance of this.

The second desideratum is that any account of pointing must be able to explain why pointing is sensitive to whether the other person has seen the object before (Liszkowski et al., 2008). And here Gómez's version of PR seems to be at a disadvantage compared to NPR. Whether an agent is attending to something may or may not be perceptually represented but it is extremely unlikely that whether the agent had seen an object before is perceptually represented (but see Apperly and Butterfill 2009 for a different way of dealing with this desideratum).

How about the Vicarious Perception version of PR? Is it more suitable to meet this challenge than Gómez's version? In order to answer this question, we need to look at the details of the experiment in Liszkowski et al (2008). In these experiments, there were two scenarios: The child observes the other agent (the experimenter), who either sees the object sliding down to a specific spot (knowledge condition) or she doesn't (ignorance condition). Then she is actively searching for the object and the finding is that the child points to the object more often in the ignorance condition. How can we explain this in the Vicarious Perception framework? One possibility would be to say this: the difference between the two scenarios is that the child does not take the

experimenter to be searching in the knowledge condition – why would she be searching for the object when she must know exactly where it is? Maybe she is just joking. In the ignorance condition, in contrast, she is performing the genuine action of searching for the object. In other words, in the ignorance condition, the child attributes the property of performing an action (the action of searching) to the experimenter, whereas in the knowledge condition she doesn't. And as vicarious perception very much depends on what action one takes the other agent to be performing, in the ignorance condition, the child sees the object as affording an action to the experimenter (the action of finding it and reaching for it), whereas in the knowledge condition, she does not see it as affording any action to the experimenter as the experimenter is not performing any bona fide action.

Conclusion

We have explored three different conceptual frameworks for describing the representational requirements of pointing: that pointing presupposes non-perceptual representations, that pointing does not presuppose any representation at all and that pointing presupposes perceptual representation. Although at least one of us endorses a version of the last one of these views (Nanay, 2013, Chapter 6), the aim here was not to argue for this account, but to consider the advantages and disadvantages of these three options. Each of these accounts have their strengths and weaknesses when it comes to explaining pointing production – they explain some aspects of pointing better than others. Our aim was to provide a road-map for the possible ways in which we can frame the representational requirements of pointing production.

ACKNOWLEDGMENTS

This work was supported by the EU FP7 CIG grant PCIG09-GA-2011-293818 and the FWO Odysseus grant G.0020.12N.

REFERENCES

- Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states?. *Psychological Review; Psychological Review*, 116(4), 953–970 .

- Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly*, 21, 205–226.
- Behne, T., Liszkowski, U., Carpenter, M., & Tomasello, M. (2011). Twelve-month-olds' comprehension and production of pointing. *British Journal of Developmental Psychology*, 30 (3), 359–375.
- de Bruin, L., Strijbos, D., & Slors, M. (2011). Early social cognition: alternatives to implicit mindreading. *Review of Philosophy and Psychology*, 2(3), 499–517.
- Bruner, J. (1983). *Child's Talk*. New York: Norton.
- Burge, T. (2010). *Origins of objectivity*. Oxford, UK: Oxford University Press.
- Butterworth, G., & Morissette, P. (1996). Onset of pointing and the acquisition of language in infancy. *Journal of Reproductive and Infant Psychology*, 14(3), 219–231.
- Campbell, J. (2002). *Reference and consciousness*, Oxford: Clarendon Press.
- Carruthers, P. & Smith, P.K. (eds.) (1996). *Theories of Theories of Mind*. Cambridge University Press.
- Cheney, D.L., & Seyfarth, R.M. (1980). Vocal recognition in free-ranging vervet monkeys. *Animal Behaviour*, 28, 362–367.
- Crane, T. (2009). Is perception a propositional attitude?. *Philosophical Quarterly*, 59(236), 452–469.
- Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition*, 107(2), 705–717.
- Currie, G., & Ravenscroft, I. (2002). *Recreative Minds: Imagination in Philosophy and Psychology*, Oxford: Clarendon Press.
- Davidson, D. (2001). *Essays on actions and events*. Oxford University Press.
- Dennett, D. C. (1989). *The Intentional Stance*. Cambridge (Mass.): MIT Press.
- Dreyfus, H. L. (2002). Intelligence Without Representation – Merleau-Ponty's critique of mental representation the relevance of phenomenology to scientific explanation. *Phenomenology and the Cognitive Sciences*, 1(4), 367–383.
- Franco, F., Perrucchini, P., & March, B. (1996). Is infant initiation of joint attention by pointing affected by type of interaction?. *Social Development*, 18, 51–76.
- Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford, UK: Oxford University Press.
- Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition*, 17 (2), 535–543.
- Gallagher, S. (2012). Interactive Coordination in Joint Attention. Axel Seeman (Ed.), *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience*, Cambridge (Mass.): MIT Press.

- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12), 493–501.
- Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naive theory of rational action. *Trends in cognitive sciences*, 7(7), 287–292.
- Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. Cambridge (Mass.): Oxford University Press.
- Goldman, A. I. (2009). Mirroring, simulating and mindreading. *Mind & Language*, 24(2), 235–252.
- Gómez, J. C. (2005). Joint attention and the notion of subject insights from apes, normal children, and children with autism. In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and Other Minds* (pp. 65–84). Oxford, UK: Oxford University Press.
- Gómez, J. C. (2007). Pointing behaviours in apes and human infants: A balanced interpretation. *Child development*, 78(3), 729–734.
- Hutto, D. D. (2008). *Narrative Folk Psychology*. Cambridge (Mass.): MIT Press.
- Kelly, S. (2010). ‘The normative nature of perceptual experience’, In: Nanay, B. (ed.) *Perceiving the World*. New York: Oxford University Press.
- Kim, I. K., & Spelke, E. S. (1999). Perception and understanding of effects of gravity and inertia on object motion. *Developmental Science*, 2(3), 339–362.
- Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others’ beliefs in human infants and adults. *Science*, 330(6012), 1830–1834.
- Leavens, D. A., Hopkins, W. D., & Bard, K. A. (2005). Understanding the Point of Chimpanzee Pointing: Epigenesis and Ecological Validity. *Current Directions in Psychological Science*, 14(4), 185–189.
- Leavens, D. A. (2012). Joint Attention: Twelve Myths. *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience*, 43.
- Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12- and 18-month-olds point to provide information for others. *Journal of Cognition and Development*. 7(2), 173–187.
- Liszkowski, U., Carpenter, M., & Tomasello, M., (2008). Twelve-month olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition*, 108 (3), 732–739.
- Liszkowski, U. (2010). Deictic and other gestures in infancy. *Acción Psicológica*, 7(2), 21–33.
- Masataka, N. (2003). From index-finger extension to index-finger pointing: Ontogenesis of pointing in preverbal infants. In S. Kita (Eds.), *Pointing:*

- Where language, culture and cognition meet* (pp. 69–84). Mahwah, NJ, USA: Lawrence Erlbaum Associates, Publishers.
- Matthen, M. (2010). How things look (and what things look that way). In: Nanay, B. (ed.) *Perceiving the World*. New York: Oxford University Press.
- Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198(4312), 75–78.
- Nanay, B. (2010). Attention and perceptual content. *Analysis* 70: 263–270
- Nanay, B. (2011a). Do we sense modalities with our sense modalities? *Ratio*, 24 (3), 299–310.
- Nanay, B. (2011b). Do we see apples as edible?. *Pacific Philosophical Quarterly*, 92(3), 305–322.
- Nanay, B. (2012a). Action-oriented Perception. *European Journal of Philosophy*, 20(3), 430–446.
- Nanay, B. (2012b). Perceptual phenomenology. *Philosophical Perspectives*, in print
- Nanay, B. (2013). *Between Perception and Action*. Oxford: Oxford University Press.
- Nichols, S., & Stich, S. P. (2003). *Mindreading: An integrated account of pretence, self-awareness, and understanding other minds*. Clarendon Press/Oxford University Press.
- Ratcliffe, M. (2007). *Rethinking Common-sense Psychology*. Palgrave Macmillan.
- Reddy, V. (2000). Coyness in early infancy. *Developmental Science*, 3(2), 186–192.
- Reddy, V. (2003). On being the object of attention: implications for self–other consciousness. *Trends in cognitive sciences*, 7(9), 397–402.
- Reddy, V. (2005). Before the ‘third element’: understanding attention to self . In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and Other Minds* 85–109
- Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, 4(8). 299–309
- Seemann, A. (2010). The Other Person in Joint Attention A Relational Approach. *Journal of Consciousness Studies*, 17(5–6), 5–6.
- Seemann, A. (2012). Joint Attention: Toward a Relational Account. Axel Seeman (Ed.), *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience*, Cambridge (Mass.): MIT Press.
- Siegel, S. (2005). The Phenomenology of Efficacy. *Philosophical Topics*, 33 (1), 265–84.
- Siegel, S. (2006). Which properties are represented in perception? In T. Gendler & J. Hawthorne (Eds.), *Perceptual Experience* (pp. 48–503). Oxford: Oxford University Press.
- Siegel, S. (2009). The visual experience of causation. *Philosophical Quarterly*, 59

- (236), 519–540.
- Sterelny, K. (1990). *The Representational Theory of Mind*. B. Blackwell.
- Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Belknap Press.
- Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: Research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry*, 42(1), 3–48.
- Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, 78, 705–722.
- Van Gelder, T. (1997). Dynamics and cognition. *Mind design {II}: Philosophy, Psychology, Artificial Intelligence*.
- Vygotsky, L. S. (1978). *Mind in Society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wellman, H. M., Cross, D., & Watson, J. (2003). Meta-analysis of theory-of-mind development: the truth about false belief. *Child development*, 72(3), 655–684.
- Zahavi, D. (2008). Simulation, projection and empathy. *Consciousness and Cognition*, 17(2), 514–522.
- Zlatev, J., Racine, T. P., Sinha, C., & Itkonen, E. (2008). What makes us human? *The Shared Mind: Perspectives on Intersubjectivity*, 12(1), 1–14.

