Being Nativist about Mindreading: More Demanding than You Might Think

Marco Mazzone (mazzonem@unict.it)
Department of Human Sciences, 32 Piazza Dante
Catania (CT) 95124 ITALY

Abstract

Two distinct theses are sometimes argued for, or against, together in the debate on early mindreading: that infants mindread, and that this occurs thanks to innate mental notions. When this is the case, the underlying assumption is that mental notions cannot be learned early in infancy, and therefore at that stage they must be either innate or not present at all. I do not intend to directly argue in favour of the opposite claim that mental notions can be learned early in infancy. My more indirect route is to argue that a much stronger form of nativism than what is ordinarily thought is required if innate mindreading abilities are to be attributed to infants. This means that, if one defends early mindreading, the choice is between acknowledging that infants can learn to mindread, on the one hand, and being committed to a rather cumbersome version of nativism on the other. My argument is based on the claim that mental notions need to subsume – and therefore generalize over – a host of behavioural rules. This argument is assessed against the nativist position of Baillargeon and colleagues and the anti-nativist position of Ruffman. Finally, I consider a possible objection to the claim that mental notions essentially consist of statistical generalizations from experience.

Keywords: mindreading; nativism; development; generalization; false belief task

Introduction

In the most recent debate about mindreading, especially after Onishi and Baillargeon (2005), a major issue is whether mental notions are a) innate and thus b) available to children very early (between the defenders of early mindreading, see for instance Southgate, Senju & Csibra, 2007; Surian, Caldi & Sperber, 2007; Song & Baillargeon, 2008; and on the opposite side, between the critics of early mindreading, Ruffman, Taumoepeau & Perkins, 2012; Heyes, 2014; Ruffman, 2014). Although a) and b) are two distinct theses, they are occasionally argued for, or against, by a single line of argument, to the point that they may not even be perceived as *two different* theses at all.

This is not completely unreasonable since if mental notions were in place as early as the first year of life or even earlier, this might suggest that they are innate; while, on the contrary, if you do not sympathize with a nativist account of mental notions, you might also be inclined to assume that these notions require a reasonably long development, possibly "influenced by cultural processes and closely tied to language acquisition" (De Bruin & Newen, 2012: 241). However, a) and b) are in principle two distinct issues, a point that is especially relevant for critics of nativism. One who is confident about the power of learning processes might acknowledge that infants mindread without conceding the existence of innate mental notions, a position that has been recently argued for by De Bruin & Newen (2012). But, as I will show, some anti-nativists such as Ruffman feel committed to argue that infants *do not* posses mental notions, so that he seems to share with nativist the assumption that mental notions cannot be learned early in infancy: they must be either innate or not present at all.

In the present paper, I do not intend to directly argue against this assumption that mental notions cannot be learned early in infancy – and therefore, in favour of powerful learning processes. My purpose is rather to frame the dispute between nativist and anti-nativist in an unusual way, based on the following claim: at a closer analysis, a much stronger form of nativism than what is ordinarily thought is required in order for innate mindreading abilities to be attributed to infants. This means that the choice is between acknowledging that infants can learn to mindread, on the one hand, and being committed to a rather cumbersome version of nativism on the other.

My line of argument is based on the issue of parsimony. Defenders of early mindreading claim that mental notions allow to get rid of the complexities inherent in behavioural explanations, since such notions generalize over a variety of state of affairs and behaviours. Now, I will claim, in a sense this is clearly true, which is the reason why the thesis of early mindreading might turn out to be correct, after all. But in another sense, the opposite is true. As the critics of early mindreading have urged, mental notions cannot be simpler representations than behavioural rules: they reduce a multiplicity to a unity but only at the price of becoming more, not less, complex, precisely because they have to subsume a multiplicity of behavioural rules. By not appreciating this fact, some nativists underestimate what they are really committed to.

I will analyse this claim with special regard to the work of Baillargeon and colleagues and, specifically, I will argue that the above problem does not affect their explanation of the infant evidence as involving mindreading: the appropriate generalizations are presumably available to children very early (thesis b) even if they are not likely to be innate (thesis a). On the other hand, anti-nativists such as Ruffman seem to reject early mindreading essentially because they want to bar the way to nativism, more specifically, to the idea that mental notions are present from the beginning in humans independently of experience. But once the distinction between thesis a) and b) is made clear, there are no reasons why Ruffman should deny that infants mindread, given their abilities for generalizations over experience.

Finally, I will consider an objection – recently raised by Perner (2014) and by Christensen & Michael (in press) amongst others – to this suggestion that mental notions are learned by simple associative generalization. The basic idea

is that there must be more to mindreading than mere associative abilities. I will argue that what this objection actually calls our attention to is the need for a crucial distinction between two ways to conceive of associative processes.

Mindreading and Parsimony

Starting from Premack & Woodruff's (1978) seminal paper, defenders of the mindreading interpretation have always appealed to mental notions as a way to avoid (what they claim to be) implausibly complex explanations based on behavioural rules. This is how Onishi & Baillargeon (2005: 257), for instance, invoked a principle of parsimony:

> it is more parsimonious to assume that infants attribute to others beliefs that can be shaped and updated by multiple sources of information than to assume that infants form an extensive series of superficial expectations linking different perceptions to different actions.

On the other hand, however, critics of early mindreading have opposed to this view and argued that the principle of parsimony is rather on their side. The argument has been put forth by Penn & Povinelli (2007) in the context of the debate about mindreading in non-human primates and then extended by Perner (2010) to the debate on infants. It is based on the assumption, which is shared by most of their opponents, that mental notions are "intervening variables" in the sense of Whiten (1996), in other words, they are *unobservables* connecting sets of observable clues and sets of behaviours expected to occur in the presence of those clues. Now, given this assumption, Penn & Povinelli (2007) observe that, for any explanation in which behaviours are indirectly inferred from clues by means of an unobservable mental state, there is always a more parsimonious explanation connecting clues and behaviours directly.

As is the case in many theoretical debates, both parties tend to dismiss hastily the opponents' argument, while I claim that they have something to learn from it.

To start with, what defenders of early mindreading call our attention to are the benefits of generalizations. As a matter of fact, Perner (2010) acknowledged that, although it is true that any mental explanation of single cases can be substituted for by a more parsimonious behavioural one, this strategy becomes less and less plausible as the number and variety of cases that are subject to experimental scrutiny grows, since one needs to postulate a new behavioural rule for each case instead of a limited number of general explanations based on mental states attribution. Many are convinced nowadays that the evidence accumulated so far is sufficient to conclude in favour of a few mental explanations instead of a host of (quite complex)[1] behavioural ones. But I want to emphasize that, if you are not sympathetic with nativism, this is not bad news in itself. Generalization is a key principle of learning from environmental regularities; as a consequence, anti-nativist should be at ease with the idea that generalizations over behavioural rules spontaneously occur as these rules grow in number.[2]

However, my main focus here is on a complementary point: what nativists can learn from the anti-nativist version of the parsimony argument.

Which Nativism for Mental Notions?

Perner (2010) made a general point that was intended against early mindreading but that lose none of its importance even if one concedes that infants mindread, after all. The point is that nativism is an explanatory strategy by which many have exempted themselves from the necessity to provide actual explanations of cognitive processes. While if you claim that something is learned you are held responsible for explaining how this may occur, who makes an appeal to innate capacities (especially when this is conjoined with modularism) is usually not required to be specific about the actual mechanisms to any comparable extent. In a sense, Perner claims, nativism (plus modularism) as a default explanatory strategy is a potential treason to the cognitive turn: instead of looking into the "black box" of behaviourists and disclosing its internal functioning, nativists tend to provide us with an entire system of black boxes. More to the point, this means that nativist explanations may appear simple just because they conceal any internal complexity. Thus, it is much too easy to claim that innate mental notions provide us with less complex explanations than the whole mess of behavioural rules: unless the internal mechanisms are disclosed this claim is ungrounded.

Perner (2010) aims to show by this, specifically, that the parsimony argument of the defenders of early mindreading is faulty, and therefore that it gives no reason to attribute mental notions to infants. On the one hand, this conclusion might overlook the fact that there is a clear sense in which mental notions are "simpler" than behavioural rules even in case the former are internally complex: as generalizations, mental notions provide a single mechanism in substitution for a plurality of unrelated rules. This sort of parsimony might turn out to be decisive for the debate on early mindreading and, as I said above, it is not something that in itself would jeopardize anti-nativism, since a tendency to form generalizations is inherent in our ability to learn from regularities. On the other hand, Perner's argument reveals its full consequences only if considered together with Penn & Povinelli's (2007) version of the parsimony argument.

The crucial fact is that mental notions can hardly be conceived of as simple entities, if they have to play the role of "intervening variables". They can play such a role only insofar as they are properly connected to the right classes of observable clues and behaviours – otherwise they could not license prediction of those behaviours from those clues – which is precisely what allows Penn and Povinelli to claim

[1] On this point, see Carruthers (2013: note 7).

[2] For a more detailed defence of this view of mental notions as generalizations, see Mazzone 2014.

that they are dispensable, in the end. If we have the right class of observable clues, the right class of behaviours and the right connections between them, we can apparently dispense with any intermediary notion. To be sure, Penn and Povinelli's argument does not take into consideration the benefits of generalization: there is a clear sense in which one single mechanism is more parsimonious than many. But the point is how such generalization is made possible in the first place. And here is where Perner's argument comes in: unless nativists specify a mechanism for this sort of generalization, one cannot say whether this mechanism is simpler than anything else.

So let us ask what kind of mechanism might be required in principle, in order to generalize over behavioural rules. I start from the side of observable clues. What we need is, first, to form classes of clues that differ from each other for features that are irrelevant to the task at hand and, second, to properly link together such distinct classes of clues to the effect that they, either disjunctively or conjunctively, allow the prediction of a certain class of behaviours. And the same is required on the side of behaviours: one needs that entire classes of distinct behaviours are linked together and properly associated, either disjunctively or conjunctively, to the set of clues they can be predicted by. When a generalization of this sort is in place, an infant can be expected to predict behaviours from observable clues in a general and flexible way, compared with a condition in which the infant just possesses a limited number of unrelated behavioural rules. It should be emphasized, however, that a mechanism of this sort is hardly simpler than a set of behavioural rules: as a matter of fact, it incorporates all the computational complexity of a set of behavioural rules and much more further, insofar as it has to compute – so to speak – a virtually infinite number of behavioural rules.

Thus, if one postulates innate mental notions in order to explain how infants go beyond – and specifically, generalize over – mere behavioural rules, then what needs to be innate is a mechanism of that sort. Simply assuming that such a mechanism is coupled (or associated in whatever way) to an innate mental notion will not do since, if the notion is innate but the mechanism is learned, then the explanation of how infants generalize lies on the learned, not the innate, component.

This rules out the idea that innate mental notions explain mindreading, if, for instance, these notions are conceived of as atomic labels, as in Fodor's (1975) theory of concepts. For example, if the notion of BELIEF is a label in a language of thought, without any internal semantic structure, then a distinct mechanism is needed in order to apply the label to the appropriate clues and predict the appropriate behaviours.[3] And if this mechanism is not innate, then there is *no innate explanation* of how infants generalize over behavioural rules.

In sum, my disjunctive conclusion is the following: since infants need complex generalizations in order to go beyond mere behavioural rules, either *such generalizations* are innate or there is no innate explanation of mindreading. In this latter case, one should consider with renewed interest a learning account of mindreading. In the former, one is committed to a much more demanding nativism than what is usually acknowledged.

Let me emphasize how my argument is different from the objection raised by Penn & Povinelli (2007) and Perner (2010). Their focus is on the fact that when explanations are based on mental states attribution, their structure must include "intervening variables" interposed between sets of clues and behaviours, which makes those explanations more complex than the corresponding behavioural rules in the precise sense that there is one intermediate step. This addition of an intermediate step has no explanatory power – in that a simpler behavioural explanation can be provided in any specific case, and therefore it is not clear why one should introduce such unnecessary complication – *unless* it turned out that infants' understanding of behaviour is general and flexible, in which case intervening variables would provide exactly the required generalizations. Thus, Penn & Povinelli (2007) and Perner (2010) aim to provide an argument against mental explanations, one which is in fact overcome in case – as is thought by many – the number and complexity of behavioural rules grow beyond reasonable limits. What I provide is instead an analysis of the notion of interposed variable showing that, even if the argument is overcome *as an argument against mental explanations*, it still has consequences that have failed to be fully acknowledged. Specifically, the argument shows that, since mental states are explanatorily required in that they provide generalizations over and above behavioural rules, what must be innate if nativism has to be right is the entire structure by which such generalizations are ensured, that is, the entire pattern of connections between sets of clues and behaviours.

In sum, by directing their argument against mental state attribution, Penn & Povinelli (2007) and Perner (2010) have probably underestimated the role of generalizations in infants' understanding of behaviour. But they have also overlooked that the argument could have been marshalled for a different purpose, that is, as directly aimed against nativism.

What are preferences, beliefs, etc.?

In order to better understand the implications of our argument, let us consider how a specific case is analysed from both the nativist perspective of Baillargeon and colleagues and the anti-nativist perspective of Ruffman. According to Baillargeon et al. (2010: 115) numerous

[3] In fact, a generalized version of Penn and Povinelli's argument might raise a serious problem for the atomistic theory of concepts as such. Conceiving of concepts as labels implies that they are wholly inert entities, with no essential role in any cognitive explanation. For them to play any such role, something else must tell the system what to do with them. But then, it is not clear that the labels are needed anymore.

experiments "contradict the notion that infants merely form associations":

> To illustrate, after watching familiarization events in which an agent repeatedly grasps object-A, infants look longer at test events if the agent now grasps object-B, but only if object-B is both present and visible to the agent during the familiarization events, so that infants have evidence that the agent prefers object-A over object-B.

In other words, infants seem to be sensitive to the agent's preferences, which are made manifest just in case there is an alternative to the object that she repeatedly grasps. In the absence of such an alternative in the familiarization stage, infants do not show expectations about whether the agent will grasp object-A or object-B, as would be expected instead if they had simply formed an association between agent and object-A. This argument against mere associations apparently dismisses the possibility that more complex, conditional associations are formed and that infants base their expectations on such more sophisticated associations. The underlying assumption seems to be that mental notions such as *preference* are not something that can (and must) be learned by associations.

The mental notion of preference is also involved in the study reported by Song & Baillargeon (2008), where in the familiarization stage infants see two toys, a doll with blue pigtails and a skunk, and an agent repeatedly reaching for the doll. Then the agent either sees or doesn't see that the toys are placed into two opaque boxes, the skunk in a box with a blue pigtail hanging from the lid ("hair" box) and the doll in a plain box. The results of the test phase are that infants look longer when the ignorant agent reaches for the plain box rather than the hair one, while they look longer when the knowledgeable agent reaches for the hair box rather than the plain one. According to Baillargeon and colleagues, this would show that infants understand the agents' behaviour by attributing mental preferences and beliefs to them: a simple association between the agent and the doll would have led instead infants to predict that agents reach the hair box in both conditions.

Ruffman (2014: 272) does not provide a detailed alternative explanation of this case, but he dismisses the conclusions of Song & Baillargeon (2008) by just saying that in the problematic condition the agent reaches the doll where she has actually seen it last, that is, in the plain box.

Clearly, there are here at play different assumptions about what can be learned by perceptual information. Ruffman assumes that perceptual information – concerning both repeated episodes in which the agent grasps the doll and a single episode in which the agent sees where the doll is placed – is sufficient for infants to form the appropriate expectations. On the other hand, defenders of early mindreading appear to assume that the relevant correlations cannot be acquired from experience.

This is a crucial point since were it not for this disagreement about what can be learned, there would be little room for divergence. On the one hand, Ruffman cannot account for the variety of results unless he makes his explanations more complex and flexible, thus acknowledging that generalizations over simple behavioural rules are made. To this extent, defenders of early mindreading seems to be right in pointing out that a collection of unrelated behavioural rules is inadequate to account for current evidence. Ruffman (2014: 276) rejects the argument that behavioural rules cannot provide a parsimonious explanation by claiming that this criticism "does not allow for any generalization on the part of the infant". But he seems not to realize that granting generalizations amounts to abandoning explanations based on isolated behavioural rules. On the other hand, when Baillargeon and colleagues appeal to mental notions they are in fact invoking generalizations over correlations between perceptual conditions and behaviours. Otherwise there would be no explanation of how perceptual conditions cause infants to attribute mental states and how this attribution in turn allows infants to foresee consequent behaviours. The fact that in verbal explanations of the evidence mental notions are referred to by simple labels is not demonstration that they are any simpler than the host of behavioural rules they generalize over. As a matter of fact, Baillargeon and colleagues do not provide any account of what mental notions could be, over and beyond such generalizations.

In practice, my suggestion is that when Song and Baillargeon (2008) explain infants' understanding in terms of the agent's *preference* for the doll and her *belief* that the doll is placed in the plain box, they are in fact implying that infants possess the appropriate generalizations over specific clues-behaviours correlations. Similarly, I suggest that Ruffman's (2014) account, despite his effort to avoid mental labels, provides an explanation of such complex cases only if the appropriate generalizations are acknowledged.

In sum, the two accounts are much more similar to each other than is usually thought. Their divergence appears to be mainly terminological, with one party labelling the required generalizations in terms of "preference", "belief", etc., and the other refusing to adopt such labels. To this extent, the description of Baillargeon and colleagues is, I suggest, preferable in that it makes clear that the host of behavioural rules is to be unified into a single mechanism, in order for infants' understanding to be complex and flexible. At the same time, the adoption of mental terms in explaining early false-belief tasks should not conceal the fact that the underlying mechanism is far from being simpler than behavioural rules.

Where the two accounts actually part company is, as I said, on the issue of what can be learned. As Onishi & Baillargeon (2005: 257) put it, "development involves primarily learning which states underlie which actions and not coming to understand that such states exist at all". In other words, the idea is that mental notions are innate, although infants have to learn the specific actions they cause, and presumably the specific clues that cause them in turn. Baillargeon and colleagues conceive of mental states

as something that exceeds what can be learned – specifically, it exceeds the system of acquired correlations between clues and actions. Mental notions are apparently thought in terms of innate concepts that form the core around which perceptual information on situations and behaviours is wrapped. However, this is very disputable for reasons I considered in the previous section. Either innate concepts already represent the complex connections they have with perceptual conditions and behaviours, and with each other as well – which is a logical possibility, but a very demanding one, and I know of no one who explicitly defends this view; or they are just idle labels, in which case what is required for their causal efficacy is the whole bunch of learning from experience that would be required anyway, were there no innate labels at all. Ruffman, on the other hand, insists on the power of learning: the first sections of Ruffman (2014) are coherently devoted to "statistical learning and pattern recognition" and "the importance of perception". From this point of view, his approach is preferable, I suggest, in that it makes clear that the appropriate generalizations are presumably not innate. However, this should not conceal the fact that, in order to ensure the needed flexibility, explanations must go beyond isolated behavioural rules and incorporate unifying generalizations that can appropriately be considered as mental, not behavioural, rules.

Ruffman (2014: 265) tries to resist this conclusion by speaking of an "*implicit* understanding of behaviour". But in a sense everyone agrees that the mental notions involved in the explanation of early mindreading are used "implicitly": they are not explicitly accessed by the linguistic system in the first place. This should be kept clearly distinct from the fact that appropriate generalizations, be they explicitly accessed or not, are needed anyway. For both these aspects – the need for flexible generalizations and their implicit use – Baillargeon and Ruffman are closer than they are ready to admit.

However, even if I were right in assuming that their main divergence concerns the nativism issue, there is another general point of disagreement between accounts of infants' mindreading that is worthy of consideration.

Mental models

In a response paper, Perner (2014: 294) has applauded Ruffman (2014) "for cautioning us against interpreting early sensitivity to others' beliefs as evidence for an innate theory of mind and for making room for learning", but at the same time he has cautioned "against [Ruffman's] claim that all infants need is to understand that people act depending on what they perceive". More precisely, Perner's suggestion is that the information about what the agent has witnessed "must be used to define a model of the world for the agent, within which the infant can reason about what the agent should do" (Perner, 2014: 295). There are two points to this suggestion. First, theory of mind research should be more concerned with the fact that we understand people as acting for reasons in a way that goes beyond detecting lawful regularities in behaviour. Second, this seems to require that we make use of the available perceptual information in a way that cannot be conceived – so to speak – as a trivial case of perceptual categorization. Specifically, Perner suggests a role for the fact that during processing infants may keep "experiential records" (in the sense of Perner & Roessler, 2010) or records of what they have registered (in the sense of Apperly & Butterfill, 2009).

In a similar vein, Christensen & Michael (in press) have warned against the view that statistical learning is all there is to infants' understanding of behaviour, and they have proposed instead a role for causal schemas and situation models.

Does this mean that any explanation of mental notions in terms of statistical associative learning is basically wrong-headed? Or associative accounts might be correct after all, on condition that we adopt a sufficiently powerful notion of associative learning? I propose to opt for the second answer. Although the issue would require a deeper discussion, it is important to sketch the general lines of a possible answer, given the role assigned here to associative generalizations.

For one example, let us consider again the task devised by Song & Baillargeon (2008). And let us suppose that a generalization can be formed such that repeated grasping of an object (here, the doll) is represented as a preference (doll vs. skunk), and that this in turn causes the expectation that the agent will reach for the object again. However, infants also need to capture the information that the agent either sees or does not see in which box the doll is put in the present circumstance, and then they must use this information appropriately. The problem is not that no regular correlations between what an agent prefers, what she sees and how she acts might be learned from perceptual experience. The point is rather that using in context information about what the agent has recently seen – and generalizing over it – requires that this perceptual information is maintained active as long as needed for it to affect further processing properly. And this seems to require more than mere associations between co-occurrence contingencies.

In fact, the term "association" is used in psychology either in a narrow or in a broad sense. In the narrow sense, psychologists call associative (versus inferential or cognitive) a process that does not involve attention, awareness or working memory, based on the detection and coding of simple co-occurrences and their recovery by mere spreading activation. Given this notion, associative processes are clearly not sufficient to ensure that all the relevant information is made available when infants have to understand actions. On the other hand, there is a broader sense in which even cognitive processes involving working memory can be said to be associative. This is what has been called the implementation sense of "associative" (Buckner, 2011; Mazzone, submitted): both associative (narrow sense) and cognitive processes have to be in the end implemented in Hebbian networks where activation spreads based on previous experience. What is characteristic of cognitive

processing is that activation is sustained by long-distance loops in these networks. Such sustained activation is also crucial in order to detect and code extended temporal contingencies. Thus, the possibility to generalize over complex actions relies on the associative ability (implementation sense) to detect statistical regularities extended in time, thanks to processes that involve sustained activation and are therefore non-associative (narrow sense).

The claim has been made that our brain is able to abstract from experience hierarchically organized representations with increasing temporal scales, with the prefrontal cortex at the top of this hierarchy (Fuster, 2001; Hari & Parkkonen, in press). At the same time, the prefrontal cortex is thought to be involved in the long-distance loops that are needed for sustained activation and working memory (Dehaene et al., 2006). Therefore, it is reasonable to think that the prefrontal cortex is where the highest schemas for action are coded, and that the activation of such schemas is crucial for the online construction of situation models (Miller & Cohen, 2001; Cooper & Shallice, 2006). In this perspective, the warning that, in the explanations of early mindreading, mental models should be considered over and beyond what can be obtained by simple associative (and statistical) learning is correct or not depending on which notion of associative learning we are considering. In a sense, mental models consist precisely in the occurrent activation of schemas that are abstracted from experience by associative generalization.

References

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

Buckner, C. (2011). Two approaches to the distinction between cognition and 'mere association'. *International Journal of Comparative Psychology , 24 (4)*, 314-348.

Carruthers, P. (2013). Mindreading in infancy. *Mind & Language, 28* (2), 141-172.

Christensen, W., & Michael, J. (in press). From two systems to a multi-systems architecture for mindreading. *New Ideas in Psychology*.

Cooper, R., & Shallice, T. (2006). Hierarchical schemas and goals in the control of sequential behavior. *Psychological Review , 113 (4)*, 887-916.

De Bruin, L. C., & Newen, A. (2012). An association account of false belief understanding. *Cognition, 123*, 240-259.

Dehaene, S., Changeux, J., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences, 10 (5)*, 204–211.

Fuster, J. (2001). The Prefrontal Cortex - An Update: Time Is of the Essence. *Neuron, 30 (2)*, 319-333.

Fodor, J. A. (1975). *The Language of Thought.* Cambridge, Massachusetts: Harvard University Press.

Hari, R., & Parkkonen, L. (in press). The brain timewise: how timing shapes and supports brain function. *Philosophical Transactions B*.

Heyes, C. (2014). False belief in infancy: a fresh look. *Developmental Science, 17*, 647-659.

Mazzone, M. (2014). Mental states as generalizations from experience: a neurocomputational hypothesis. *Philosophical Explorations, 17*, 223-240.

Mazzone, M. (submitted). What kind of associative and inferential processes? A response to Rubio-Fernández.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience , 24*, 67-202.

Onishi, K., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*, 255-258.

Penn, D., & Povinelli, D. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical Transactions of the Royal Society B - Biological Sciences, 362*, 731-744.

Perner, J. (2010). Who took the cog out of cognitive science? Mentalism in an era of anti-cognitivism. In P. Frensch, & R. Schwarzer (Eds.), *Cognition and neuropsychology: International perspectives on psychological science* (Vol. 1, p. 241-61). New York: Psychology Press.

Perner, J. (2014). Commentary on Ted Ruffman's "Belief or not belief:…". *Developmental Review, 34*, 294-299.

Perner, J., & Roessler, J. (2010). Teleology and causal reasoning in children's theory of mind. In J. Aguilar, & A. Buckareff (Eds.), *Causing human action: New perspectives on the causal theory of action* (pp. 199–228). Cambridge, MA: Bradford Book, The MIT Press.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1* (4), 515-26.

Ruffman, T. (2014). To belief or not belief: Children's theory of mind. *Developmental review, 34*, 265-293.

Ruffman, T., Taumoepeau, M., & Perkins, C. (2012). Statistical learning as a basis for social understanding in children. *The British Psychological Society, 30*, 87-104.

Song, H., & Baillargeon, R. (2008). Infants' reasoning about others' false perceptions. *Developmental Psychology, 44*, 1789-1795.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science, 18*, 587-592.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-olds infants. *Psychological Science, 18*, 580-586.

Whiten, A. (1996). When does smart behavior-reading become mind-reading? In P. Carruthers, & P. Smith (Eds.), *Theories of theories of mind* (p. 277-292). Cambridge: Cambridge University Press