What Do False-Belief Tests Show?

Pierre Jacob

Review of Philosophy and Psychology

ISSN 1878-5158
Volume 11
Number 1

Rev.Phil.Psych. (2020) 11:1-20
DOI 10.1007/s13164-019-00442-z

Springer

Review of Philosophy and Psychology (2020) 11:1–20
https://doi.org/10.1007/s13164-019-00442-z

What Do False-Belief Tests Show?

Pierre Jacob [1]

Published online: 10 July 2019

Abstract

In a (2018) paper published in *Psychological Review*, Tyler Burge has offered a unified non-mentalistic account of a wide range of social cognitive developmental findings. His proposal is that far from attributing mental states (e.g. beliefs), young children attribute to humans the same kind of internal generic states of sensory registration that biologists attribute to e.g. snails and ticks. Burge's proposal deserves close attention: it is especially challenging because it departs from both the mentalistic and all the non-mentalistic accounts of the data so far. Moreover Burge has been one of the leading philosophers of mind of the past 40 years and some of his writings on the objectivity of perception display a deep understanding of the relevance of science for sharpening our understanding of the mind. After taking a close look at the developmental evidence, in particular at false-belief studies, I argue that Burge's (*Psychological Review*, 125(3), 409–434, 2018) account faces severe obstacles. To give one telling example: if young children can only attribute to others sensory registrations, then it is hard to explain the evidence showing that they respond differently to an agent's ignorance and to her false belief.

1 Introduction

In a recent (2018) paper published in Psychological Review, the philosopher Tyler Burge has provided a bold, original and unified non-mentalistic interpretation of a wide range of developmental findings which have been construed by a large majority of experimental psychologists as evidence of mental state attribution by young human children and even preverbal infants. Burge's contention is that all the relevant findings, including the results of false-belief tests, can be accommodated without crediting young children, let alone preverbal infants, with the mindreading capacity to attribute mental states to (self and) others. All that is required, according to Burge, is to credit young children and infants with a non-mentalistic attribution scheme, which he calls the generic action-sensing attribution scheme.

✉ Pierre Jacob
jacobpiotr11@gmail.com

[1] Institut Jean Nicod Paris, Paris, France

Burge's (2018) proposal deserves close attention for three related reasons. First, as a reviewer for this paper has crisply put it, Burge's proposal is "a challenge to pretty much the entire field of developmental social psychology." Secondly, Burge is one of the most significant philosophers of mind and language of the past forty years or so. Finally, as his (2010) monumental book on the Origins of Objectivity demonstrates, he has a deep understanding of the mutual relevance of scientific investigation and philosophy for sharpening our understanding of the human mind.

What makes Burge's non-mentalistic attribution scheme challenging for the field is that it dispenses with two basic assumptions which are widely accepted by the majority of experimental psychologists. First of all, Burge does not take success on verbal (often called "explicit") false-belief tests, in which participants are directly asked by the experimenter to predict the likely action of a mistaken agent, as reliable evidence of mindreading (or theory of mind) capacities.[1] Not all experimental psychologists agree that only success on verbal false-belief tests should be regarded as reliable evidence of false-belief attribution. But, unlike Burge, most agree that it is sufficient evidence.[2]

Some developmental psychologists take findings based on non-verbal (often called "implicit") false-belief tests (which use infants' looking behavior or brain responses as dependent measures) as reliable evidence of false-belief attribution in human infancy (cf. e.g. Leslie 1994; Baillargeon et al. 2010).[3] Others have argued that only success on verbal false-belief tests is reliable evidence of false-belief attribution (cf. e.g. Perner and Ruffman 2005; Perner and Roessler 2012; Heyes 2014a, b; Povinelli and Vonk 2004). A few psychologists do not take success on verbal false-belief tests as sufficient evidence of false-belief attribution.[4]

Burge's own view is that, whether verbal or non-verbal, "false-belief tests not only show nothing about attribution of belief. They do not in themselves show attribution of any mental states" (Burge 2018, p. 418). In short, most developmental psychologists agree that there is a puzzle about false-belief understanding in early human childhood, but disagree about its resolution. Unlike most developmental psychologists, Burge argues that there is no puzzle about false-belief understanding in early human childhood on the grounds that false-belief tests are not appropriate measures of false-belief understanding.

Secondly, Burge's non-mentalistic attribution scheme departs from all other non-mentalistic accounts so far. Some, but not all, experimental psychologists are willing to take the results of non-verbal false-belief tests at face value as evidence of false-belief

[1] There is reliable evidence that most 4,5 year-olds succeed on verbal false-belief tasks about an object's location (cf. Wellman et al. 2001).

[2] As Bloom and German (2000) were the first to argue, the capacity for false-belief attribution cannot be sufficient for success on verbal false-belief tasks, because success also requires understanding the question asked by the experimenter. If they are right, then conversely failure on verbal false-belief tasks does not demonstrate the inability to attribute false beliefs to others.

[3] Some philosophers also endorse this position, cf. Carruthers (2013).

[4] Fabricius et al. (2010) have reported an intriguing pattern whereby most of the children (between 5 and 6 years of age), who succeed on verbal false-belief tasks about an object's location, oddly fail on verbal true-belief tasks. They take these findings as evidence that these children cannot attribute genuine beliefs to others. In these true-belief tasks about an object's location, the agent's belief was true, but the agent did not know the object's location in the philosophical epistemological sense of 'know', because before being placed back where the agent had put it before leaving the scene, the toy was temporarily displaced in the agent's absence (cf. Oktay-Gür and Rakoczy 2017 for detailed discussion and for a pragmatic account of the failures in true-belief conditions). Burge does not discuss these findings which might lend support to his own point of view.

attribution in human infancy. Psychologists who are unwilling to do so have offered three basic kinds of non-mentalistic accounts and one intermediate account of infants' responses in implicit false-belief tests.

They do not take infants' looking behavior to reflect their expectations about an agent's mental states, but instead about one or another of the following three fundamental non-mentalistic features of the stimuli: the agent's observable behavior (Povinelli and Eddy 1996; Penn and Povinelli 2007); the ternary association between the agent, the object and its location (Perner and Ruffman 2005); or the low-level, non-social, perceptual features (e.g. colors, shapes and motions) exhibited by the stimuli (Heyes 2014a, b). Recently, advocates of the two-systems approach to mindreading have explored a middle ground between fully mentalistic and non-mentalistic accounts, which they construe as minimally mentalistic: while human adults have full-blown (or so-called "explicit") mindreading capacities, human infants (and non-human apes) have minimal (or so-called "implicit") mindreading capacities (Apperly and Butterfill 2009; Butterfill and Apperly 2013).

Burge's (2018) non-mentalistic attribution scheme departs from all four of these proposals in that it enables young children and infants to make two sorts of spontaneous non-mentalistic attribution to the agent of a goal-directed action. First, it enables them to attribute the executed action (not mere behavior or bodily movements) to the agent and to recognize the action's target.[5] Secondly, it also enables them to attribute to the agent what Burge calls internal but non-mental conative and sensory states, none of which are genuinely representational states.

In short, unlike any of the three prevalent non-mentalistic accounts, Burge's attribution scheme grants infants the capacity to ascribe interlocking generic non-mental internal states to an agent. Unlike Apperly and Butterfill's middle-ground two-systems approach, Burge's attribution scheme is presented as entirely non-mentalistic. On Apperly and Butterfill's two-systems approach, the minimal mindreading system (available to human infants) does not attribute to others beliefs (i.e. mental states with propositional contents or aspectuality), but registrations, which Apperly and Butterfill take to be relational mental states. By contrast, Burge's non-mentalistic scheme underlies the attribution of generic internal states that need not be mental. While Burge seems to accept the mentalistic criticisms of each of the three main non-mentalistic proposals, his main burden is to argue in favor of his own non-mentalistic attribution scheme against the mentalistic alternative (including the minimal mindreading proposal of advocates of the two-systems approach to mindreading).

To begin with, Burge (2018, p. 424) argues that there is no room for the middle-ground position that the two-systems approach wishes to occupy, because there is no room for minimal mindreading. As he puts it, "there is no such thing as a purely relational mental state that "connects" to the environment. All mental states that "connect" to the environment do so by representing it in a specific way. No science, certainly not perceptual psychology, invokes such supposed non-representational, relational mental states." There are no purely relational mental states because all relevant mental states are representational and all representational mental states are

[5] Following his (2010) book, Burge (2018, p. 410) uses "attribution" in an unusually liberal way (at least among psychologists): "Not only individuals, but states, attribute things. Perceptual states attribute properties and relations to particular entities."

aspectual. What sets genuine representational mental states apart from internal generic non-mental states is precisely that the latter lack aspectuality. To fully understand Burge's appeal to non-mental generic internal states, it is necessary first to spell out in some detail his representationalist and anti-individualist account of perception.

2 The Perceptual Level of the Representational Mind

It is impossible to do justice to Burge's (2018) non-mentalistic attribution scheme without taking into account two of his (2010) fundamental distinctions, the first of which is between perception and sensory registration. The second is the distinction between perception and propositional thought.

Burge (2018) makes a three-tiered use of his distinction between perceptual representations and states of sensory registration. (i) In his paper, Burge takes for granted his (2010) sharp criticism of the deflationary attempts by naturalistically minded philosophers, whose goal is to construe representations as internal states whose function is to carry or register information. (ii) He further points out that biologists attribute non-mental states of sensory registration to such lower animals as snails and ticks for the purpose of explaining their non-intentional actions. (iii) Finally, his fundamental thesis is that far from attributing genuine mental states (e.g. beliefs and perceptual representations), infants and young human children also explain the intentional actions of humans by attributing to them generic non-mental states of sensory registration.

Another fundamental distinction drawn by Burge (2010) is between thought (or propositional attitudes) and perception. Unlike many influential developmental psychologists (e.g. Spelke 1988 and Carey 2009), Burge (2010, chapter 10) does not think that the developmental evidence (from the study of early naïve physics and early numerical cognition) shows that human infants have a capacity to form propositional thoughts about particular objects in their physical environment. However, unlike some influential philosophers (e.g. Davidson 1982), Burge sees no reason to deny that preverbal human infants and many non-human animals entertain genuine mental representations of their physical environment. On his view, human infants form perceptual representations of their environment — in the technical sense that Burge (2010) gives to perceptual representations.

As Burge (2014) sees it, the representational mind has several levels and perception is where it begins. Perception is fully representational: it is the most primitive type of representational state. It is a natural psychological kind, recognized in a mature science, i.e. perceptual psychology. According to Burge's (2010) anti-individualist representationalist interpretation, perception provides humans and many non-human animals as well with objective and sensory representations of their physical environment.[6] Anti-individualist representationalism stands in sharp contrast with two main influential alternative standpoints: individual representationalism and deflationary conceptions of

[6] According to Burge's anti-individualist standpoint, the contents of an individual's perceptual representations depend on the individual's non-social environment. In earlier work, Burge (1979) has also argued for antiindividualism (also called 'social externalism') with respect to an individual's beliefs and other propositional attitudes by highlighting the contribution of the individual's social environment (what members of her community think and say) to the contents of her beliefs and other mental states.

representation. Burge argues that the former "over-intellectualizes" the conditions for objectivity and the latter "debase" the conditions for representation.

What is characteristic of all versions of individual representationalism is the assumption that an individual cannot form an objective representation of her physical environment unless the individual can represent conditions of objectivity. According to so-called "first-family individual representationalism," the individual must have the inferential capacities to build up representations of objective particulars from more fundamental subjective sensory experiences (e.g. representations of sense-data). According to so-called "second-family individual representationalism," the individual must have some general higher-order conception or concept of objectivity. According to the more recent deflationary tradition in naturalistic (i.e. reductionist) philosophy of mind, it is sufficient for a system to represent some environmental property that this system has the biological function to produce states that stand in some statistical informational law-like correlation with the relevant environmental property.

On Burge's anti-individualist representationalist approach, perceptual states in either humans or non-human animals are objective sensory representations of an organism's physical environment. All genuine representations have so-called veridicality conditions, which can be of two kinds: while propositional thoughts have truth-conditions, perceptual representations have accuracy conditions. The main representational task of a genuine perceptual process is to resolve or to overcome the notorious under-determination problem: its inputs are proximal (e.g. retinal) stimulations; its output is a representation of the specific distal environmental cause of the proximal stimulations; but the information at the proximal level is compatible with an indefinite set of different distal environmental causes. The hallmark of the objectivity of perception (e.g. vision) is the competence known as the perceptual constancies, namely the perceptual capacity to keep track of a particular object's invariant (or constant) attributes (e.g. size, shape, texture or color) under a wide variety of proximal (e.g. retinal) stimulations.

As I said, from Burge's anti-individualist representationalist standpoint, individual representationalism wrongly over-intellectualizes the conditions for objectivity and the more recent deflationary trend in naturalistic philosophy of mind wrongly demeans representation. In order to show what is wrong with both trends, Burge appeals to the authority of scientific explanations in perceptual psychology.

On the one hand, the standard scientific resolution of the under-determination problem faced by any perceptual process shows that the intellectualist higher-order conditions on objectivity posited by individual representationalism are dispensable. While proximal (e.g. retinal) stimulations are mere inputs to the perceptual (e.g. visual) process, the perceptual output is an objective (i.e. genuine) and sensory representation of the distal cause of the proximal stimulations. The objectivity of the perceptual representations of the distal causes of the proximal stimuli is built into the scientific resolution of the under-determination problem.

On the other hand, what is wrong with the deflationary approach to representation, according to Burge, is that by conflating the explanatory goals and conceptual resources of perceptual psychology with those of the biological sciences, it fails to account for the gap between carrying information and representing. One of the fundamental goals of the recent naturalistic tradition in the philosophy of mind is to offer a reductionist account of intentionality (or mental representation) in terms acceptable by the natural (i.e. the physical and especially the biological) sciences.

Some advocates of the reductionist program in the philosophy of mind start with the notion of information construed as the converse of a nomic (law-like) or statistical dependency: a signal carries information about some environmental property F if and only if the former statistically depends on (or nomically covaries with) the latter. Carrying information is widely recognized to fall short of representing since the latter is, but the former is not, compatible with misrepresenting. A common way to fill the gap between information and representation is to appeal to biological function along the following line: if it is the biological function of system S to produce states whose tokens carry information about environmental property F, then S represents F (cf. Dretske 1988, 1995). Other reductionist proposals simply bypass information and purport to account for the possibility of misrepresentation by directly appealing to biological functions (e.g. Millikan 2004).

While deflationist reductionists widely recognize the gap between carrying information and representing, what they fail to recognize, according to Burge, is the gap between representational error (or misrepresentation) and failure of biological function. On Burge's (2010) view (chapter 8), it is constitutive of representations that they have veridicality conditions: propositional thoughts have truth-conditions (i.e. they can be either true or false); perceptual representations have accuracy conditions (i.e. they can be either accurate or inaccurate). Burge argues that it is a deep mistake to identify representational success (truth or accuracy) with fulfillment of biological function and representational failure (falsity or inaccuracy) with failure of biological function. Unlike representational success, fulfillment of a biological function is practical success. Unlike representational failure, failure of a biological function is practical failure. Thus, perception has a biological practical function: it contributes to an agent's successful actions, which in turn contributes to the agent's biological fitness. Perception has also a different representational function, which is to enable an agent to accurately represent particulars in his physical environment and to accurately attribute properties that they instantiate so as to group together the particulars that fall under the same types.[7]

3 Burge's Non-mentalistic Attribution Scheme

Consistent with his view of perception as the most primitive level of the representational mind, Burge (2018) takes human infants and many non-human animals (e.g. non-human apes), unlike plants, bacteria, and many other non-human animals (e.g. ticks and snails), to be able to entertain objective perceptual representations of their physical environment in virtue of their perceptual capacities known as perceptual constancies. As Burge (2010, p. 274) has put it, these are "capacities systematically to represent a given particular or attribute as the same despite significant variations in proximal stimulations."

Although Burge takes infants and non-human apes to have the capacity to form first-order perceptual representations of their physical environment, he takes them to be

[7] Granting Burge his thesis that non-mental states lack representational functions, a deep question is: what gives mental states their representational functions? Burge seems to think that mental states do not derive their representational functions from evolution by natural selection. One alternative possibility highlighted by Dretske (1988), but not, so far as I can see, by Burge is that the representational functions of mental states rests on ontogenetic learning.

unable to attribute any mental states to others, including the kind of perceptual states that they themselves enjoy. Arguably, the fact that Burge takes infants to be unable to attribute mental states to (self and) others follows from the fact that Burge takes infants to be unable to form propositional thoughts. If infants lack beliefs, then they cannot form beliefs about beliefs or any other mental states. Instead, Burge credits infants and non-human apes with the non-mentalistic generic action-sensing attribution scheme.

Burge's (2018) overarching goal is to argue that his non-mentalistic generic action-sensing attributing scheme is the best current explanation so far of the developmental findings, which some psychologists have taken as evidence of false-belief attribution by preverbal infants. The first of his two subordinate goals is to establish the scientific credentials of his non-mentalistic scheme and to show that it avoids the pitfalls of the three competing non-mentalistic accounts of the developmental data (behavior-reading, associationism, low-level non-social perceptual features). Since Burge agrees with the mentalistic criticisms of the other three non-mentalistic alternatives, his second subordinate goal is to argue that his hypothesis that infants apply his generic action-sensing attribution scheme is more parsimonious (or leaner) than its mentalistic alternative, including the two-systems approach to mindreading (cf. section 5 of the present paper).

As made clear in the first section, Burge (2010) takes it as a major metascientific consideration in favor of regarding perceptual states as having veridicality (i.e. accuracy) conditions that a stable science, i.e. perceptual psychology, makes systematic reference to perceptual states as having accuracy conditions in its law-like explanations. Similarly, Burge (2018) takes it as a major metascientific desideratum to establish that his action-sensing attribution scheme is also part of stable biological scientific explanations.

The scheme rests on a crucial hierarchical distinction between generic (or superordinate) and specific (or subordinate) internal states attributable to an agent. Burge takes all of an agent's generic internal states to have a mental and a non-mental sub-species. Mental states can either be conscious or representational.[8] Both representational mental states and non-mental states can be either conative (i.e. motivational) or sensory (i.e. informational). From a scientific biological standpoint, many more animal actions are non-intentional actions caused or guided by generic internal non-mental states than intentional actions caused by mental representational states. Burge recognizes that some human intentional actions, including some intentional actions performed by human infants, and also some performed by non-human animals (in particular apes), are caused by the agent's first-order mental representations. But actions performed by bacteria, ticks or snails are non-intentional actions guided by generic internal non-mental states, not by mental states.

So internal generic states stand to mental states as genus to species or as species to sub-species. Burge's strong claim is that any sensory or conative mental state must have a non-mental internal counterpart. A non-mental conative state (e.g. a thrive) has the biological function to provide energy to cause an agent's action and thereby the realization of the action's target. But unlike mental conative states (intentions and desires), a non-mental conative state lacks representational content. Perception is a mental species of the genus sensing. A non-mental sensory (or sensing) state has the biological function to carry information about (or to statistically covary with) environmental features. But unlike mental perceptual states, a non-mental sensing state lacks accuracy conditions and representational content. So both conative states and sensory

[8] I'll ignore conscious states, which are irrelevant to the present discussion.

non-mental states have biological practical functions. A non-mental conative state shares its practical biological function with the action that it causes. The biological practical function of an agent's non-mental sensory state is to contribute to the success of the agent's action by carrying information about (or covarying with some) relevant environmental property. But since a non-mental sensory state lacks representational content or accuracy conditions, it also lacks a representational function.

A non-mental state of sensory registration in Burge's (2018) sense is just a non-representational information-carrying state (that stands in a statistical or nomic covariation relation with e.g. environmental property F), which many advocates of the deflationary approach wrongly took, according to Burge, as a basis for a naturalistic reduction of the capacity to represent F. Burge (2010) argued that all such attempts are bound to fail because they wrongly confuse representational and biological functions: biological functions can contribute to the practical success or failure of an agent's action (and fitness), but not to the veridicality conditions of a representation. In short, unlike mental states, non-mental internal states lack aspectuality (cf. section 4).

Burge (2018), however, further argues that sensory registration is part of a non-mentalistic biological scientific scheme of explanation of a wide range of non-human animal and human actions. According to this biological scheme, a snail's target-oriented action of crawling toward and eating a leaf is caused by a generic internal non-mental conative state. Bacteria sense light; mollusks sense predators in virtue of sensing light; ticks sense human arms in virtue of sensing their heat. Their generic internal sensory states carry information about, but do not represent, respectively light and heat. Similarly, a human agent may be said to automatically remove her hand from a hot surface in order to protect her hand from the heat and because she sensed the heat. Thus, Burge convincingly shows that the action-sensing scheme belongs to mature biological explanations of the non-intentional actions of many nonhuman animals and human agents as well. The relevant question for Burge's purpose, however, is: what is the evidence that the action-sensing scheme used by biologists to account for the non-intentional actions of many non-human animals is available early in human infancy to account for human intentional actions?

4 Burge's Non-mentalistic Alternative to False-Belief Attribution

In his paper, Burge discusses developmental findings, some of which (reported in particular by Csibra, Gergely and colleagues, cf. Geregely & Csibra 2003) uncontroversially show that infants take a non-mentalistic teleological stance with respect to an agent's goal-directed action. He also discusses other developmental findings (first reported by Woodward 1998 and later by Luo and Baillargeon 2005 and Luo 2010, 2011) that show that infants can ascribe a mental state of preference for one of a pair of targets to an agent. It is an open question, I think, whether Burge's non-mentalistic action-sensing scheme can provide a non-mental conative counterpart to the mental conative state of preference and especially whether it can match the tripartite mentalistic distinction between an agent's intention, her desire and her preference.

In this section, I wish to focus on Burge's deflationary non-mentalistic claim that none of the existing false-belief tests should be taken as evidence of belief-attribution in human childhood, let alone in human infancy. He argues that all false-belief tests have

the structure of his example of a tick's failed action, which can be, and from a scientific point of view, ought to be, explained by his non-mentalistic action-sensing attribution scheme (Burge 2018, p. 420).

Suppose that a tick is crawling towards a human arm whose presence it is sensing in virtue of sensing the heat emanating from it. Suppose that the tick's displacement is blocked and the human arm moved out of the tick's sensory range. Consistent with what is known of the biological process of sensing heat in ticks, suppose that if and when the tick is allowed to move again, its crawling is guided by what Burge calls "a state of sensory retention" of the former non-actual location of the source of heat, i.e. the former location of the arm. So described, the tick's failed action does not involve any misrepresentation, let alone any false belief. This is an instance of a failed action without any misrepresentation, let alone false belief: the tick is expected to miss its target because its state of sensory retention is out of sync with the location of the tick's target object. According to Burge (2018, p. 420), "the case has the structure of a false belief test."

Assuming that a state of sensory retention in the absence of a target is a non-mental state, one can agree with Burge that a biologist (or a biologically informed adult) could provide the above non-mentalistic biological account of the tick's failed non-intentional action and refrain from construing the tick's action as an intentional action guided by mental states.[9] But what is the evidence that this non-mentalistic action-sensing attribution scheme is what enables infants to make sense of human intentional failed actions? Consider what scientific background knowledge is required for recognizing the tick's non-intentional action as a failed action to begin with. One should expect the tick's crawling action to be guided by the tick's goal of sucking a human arm's blood and its capacity to sense the heat emanating from the human arm. One could only expect the tick to crawl to the arm's non-actual location (and fail to achieve its goal) if one knew that the tick's sensory retention of heat is out of sync with the human arm's known actual location. Preverbal human infants are unlikely to understand the tick's failed action as a failed non-intentional action at all.

We know, however, from a growing list of findings based on non-verbal false-belief tests (since Onishi and Baillargeon's 2005 seminal study) that infants who know the location of a toy expect an agent who wants to find it to look for it, not where they know it to be, but where she expects it to be. Moreover, scientific evidence based on brain imaging (functional near-infrared spectroscopy or fNIRS) in human adults has recently been taken to show that activity in the temporal-parietal junction in the right hemisphere (rTPJ) underlies spontaneous or perhaps automatic mindreading activity (including false-belief attribution) when adults are engaged in "free-viewing of social videos" (cf. Hyde et al. 2015). Further evidence based on fNIRS in human infants shows that the very same brain areas, which are active when adults watch false-belief scenarios played on videos and are known to be active in adults' false-belief attribution, are also active in 7-month-old infants when they watch the same videos as the adults (cf. Hyde et al. 2018). With this neuroscientific evidence in the background, I now wish to highlight three behavioral experimental studies where it seems to me unlikely that Burge's non-mentalistic scheme has the full resources to account for basic findings

[9] Burge's assumption here seems to be that sensory retention in the absence of a target can be construed as a non-mental counterpart to a mental state of believing or remembering.

based on false-belief tests. I will argue in particular that the non-mentalistic scheme faces three challenges: accounting for true-belief conditions; accounting for helping false-belief tests; and accounting for the demonstration that infants are sensitive to the difference between an agent's false belief and an agent's ignorance.

4.1 The Challenge of True-Belief Conditions

False-belief tests include a false-belief condition and a true-belief condition (and sometimes an ignorance condition). A proper interpretation of such false-belief tests is expected to coherently account for both the false-belief and the true-belief conditions (and also for the ignorance condition when there is one). To begin with, consider Burge's (2018, pp. 420–421) application of his non-mentalistic action-sensing attribution scheme to findings from a "false-perception" study by Song and Baillargeon (2008). Infants were initially given evidence that an agent had a preference for a doll with blue hair over a skunk. Then the infants saw the skunk being placed in an opaque box with a visible tuft of blue hair attached to it, while the doll was being placed in a plain opaque box, in either the agent's presence (true-belief condition) or the agent's absence (false-belief condition). Finally, the infants saw the agent reach for one box or the other. In the test events of the true-belief condition, 14.5-month-olds looked reliably longer when the agent reached for the box with the tuft of blue hair than for the plain box. The converse pattern emerged in the test events of the false-belief condition.

Burge (2018) argues that all that is needed to account for the results in the false-belief condition is that infants track the agent's target, i.e. the doll with blue hair, by sensing one of its significant properties, i.e. the tuft of blue hair attached to the box (that contains the skunk). In his own words (pp. 420–421) "the evidence is explained by taking the infant to attribute a generic capacity to chain sensing of the doll with sensing of blueness."

Although the findings in the false-belief condition are clearly open to this possible non-mentalistic interpretation, what is less clear is how this non-mentalistic interpretation can make full sense of the contrast between the results of the false-belief condition and the true-belief condition. In the test events of both the false- and the true-belief conditions, the infants see the agent confronted to a pair of boxes, one of which has a visible tuft of blue hair attached to it. On Burge's interpretation, infants in the true-belief condition should also presumably attribute to the agent the generic capacity to track the doll with blue hair by sensing the property of blue (hair) instantiated by the tuft of blue hair attached to the box containing the skunk. As a result, in the true-belief condition, infants should be torn between attributing to the agent the knowledge (or, as Burge would say, the sensory retention of the information) that the doll with blue hair is hidden in the plain box and attributing to her the non-mental capacity to track the doll with blue hair by sensing the currently visible tuft of blue hair attached to the other box. It is hard to reconcile this expected tension with the clear-cut findings in the true-belief condition.

4.2 The Challenge of Helping False-Belief Tests

A second challenge for Burge comes from the evidence that toddlers provide appropriate help to a mistaken agent. For example, 24- and 18-month-olds have been shown by Knudsen and Liszkowski (2012) to spontaneously point to an object's location

(which they know), for the benefit of an agent, if but only if the agent's goal is to retrieve the object and the agent has a false belief about its location, not if the agent either knows the object's location or her goal is not to retrieve it.[10] 18-month-olds have also been shown by Buttelmann et al. (2009) to help a mistaken agent find her toy by opening the box that contains the toy if the toy was placed there in the agent's absence (false-belief condition), but not in if the toy was placed there in the agent's presence (true-belief condition).

Burge (2018, p. 421) argues that this particular case can be accounted for by his non-mentalistic action-sensing scheme "together with the assumption that the child wants to help the actor reach the actor's target." However, as argued by Carruthers (forthcoming), the question is: what grounds this assumption? Why should toddlers feel motivated to help a mistaken agent if they merely ascribed to her the non-mental conative state of seeking to achieve an environmental goal-state characterized along the lines of the teleological stance? It seems as if toddlers themselves could only be motivated to help a mistaken agent if they recognized that whether the agent's mental state of desire to find her toy is fulfilled or frustrated makes what I take to be a mental difference to the agent's subsequent emotional (or affective) state. Note that in the absence of any emotional behavioral expression, toddlers must anticipate the mistaken agent's likely emotional response that would arise would her desire to find her toy be frustrated. Burge (2018) argues that all affective and emotional mental states have non-mental counterparts. But as I understand it, such non-mental emotional states could only be ascribed by a toddler to an agent on the basis of the execution, not in anticipation, of the agent's relevant action. In the helping study by Buttelmann and colleagues under discussion, in order to feel motivated to help, the toddlers must ascribe to the mistaken agent an emotional response which would be caused by her discovery that the box that she believes not to be empty is empty.

4.3 The Challenge of the Dissociation Between False-Belief and Ignorance Attribution

A third important empirical challenge for Burge's claim that the biological non-mentalistic account of the tick's failed action is a good model of infants' responses to false-belief tests is the evidence showing that they do not respond in the same way to an agent's ignorance and to an agent's false belief. In a famous false-belief study mentioned but not discussed in detail by Burge, Scott and Baillargeon (2009) have investigated three distinct conditions: a false-belief, a true-belief, and an ignorance condition, involving a pair of penguins, one of which was a one-piece penguin and the other a two-piece penguin that could be assembled. When assembled, the two-piece penguin looked indistinguishable from the one-piece penguin. In the familiarization trials, the infants saw an agent repeatedly place her key within the disassembled penguin and assemble it. In the first stage of the test trials, a human hand assembled the disassembled penguin, placed it into a transparent box and placed the one-piece penguin into an opaque box, either in the agent's absence (false-belief condition) or in the agent's presence (true-belief condition). In the ignorance condition, everything was

[10] Knudsen and Liszkowski (2012) also rule out the possibility that infants attributed ignorance, rather than false belief, to the agent.

similar to the false-belief condition except that the two boxes were identical. In the test events of the ignorance condition, half the infants saw two transparent boxes; half saw two opaque boxes. In the final stage of all test trials, the agent opened either one or the other box, while ostensibly holding her key.

Scott and Baillargeon (2009) found a different pattern of responses in all three conditions: infants' looking times were reversed in the false- and the true-belief conditions, but the infants did not look reliably longer at either of the two events in the ignorance condition. What makes these findings challenging for Burge is that in both the false-belief and the ignorance conditions the agent was unable to sense or track which penguin was placed in which box (since she was absent). But nonetheless the infants did not respond to the ignorance condition and the false-belief condition in the same way.

The mentalistic distinction between ascribing ignorance and ascribing false belief to an agent is crucial for human social cognition: each ascription warrants different behavioral predictions. Ascribing ignorance to an agent about whether a toy is in the green or in the yellow box warrants the prediction that if the agent wants to find the toy, then she is equally likely to search for it in either box. Ascribing to the agent the false belief that the toy is in the yellow box warrants the prediction that it is highly likely that she will search for the toy in the yellow box. On Burge's non-mentalistic action-sensing attribution scheme, an agent is either provided with the opportunity for sensing a property or not. So the challenge for Burge is whether there is room within his non-mentalistic scheme for the distinction between two distinct ways of failing to sense, track or register, one and the same relevant piece of information that could match the mentalistic distinction between ignorance and false belief.[11]

5 Burge on False-Belief Tasks About Object-Identity

As emphasized by philosophers since Frege's (1892) well-known puzzle, Mara may believe that Cicero was a Roman orator while disbelieving that Tully was a Roman orator even though Tully was Cicero. Mara may hold one belief, not the other, if e.g. she does not know that 'Cicero' and 'Tully' are coreferential. In other words, an agent may hold two conflicting beliefs about a single state of affairs consisting of the instantiation of a single property by one and the same individual. This property of beliefs (and other propositional attitudes) is known as the aspectuality of beliefs. To understand the aspectuality of beliefs is to understand that one and the same fact can be represented in different ways and that the way an agent represents the relevant fact matters to the content of the agent's belief.[12]

For the past fifteen years or so developmental psychologists have investigated in some details the capacity of young children to understand the aspectuality of beliefs. In particular, a major claim of the two-systems approach to mindreading championed by Apperly and Butterfill (2009) is that understanding the aspectuality of beliefs is a

[11] Further significant evidence that toddlers (and even infants) construe differently an agent's state of ignorance from an agent's false belief comes from He et al. (2011) and Knudsen and Liszkowski (2012).

[12] The aspectuality of beliefs explains the intensionality (or referential opacity) of belief-attributions, exhibited by the fact that replacement of 'Cicero' by the coreferential name 'Tully' does not preserve the truth of the ascription 'Mara believes that Cicero was a Roman orator'.

"signature limit" of infants' minimal mindreading capacities and a hallmark of genuine full-blown mindreading (available to older children and adults). Much of the developmental psychology of mindreading rests on the assumption that the most perspicuous way to tell whether children understand the aspectuality of beliefs is to test their capacity to attribute false beliefs about object-identity (not merely false beliefs about an object's location) by means of false-belief tests about object-identity. In such tests, young participants who know that a single object (or individual) has two different aspects or properties are requested to attribute to a mistaken agent the false belief that two distinct objects (or individuals) exemplify the two aspects or properties.[13]

For example, children are provided with the evidence that an agent enjoys playing with a bunny-toy, which the agent places into one of a pair of boxes (e.g. the green box) before she departs. In the agent's absence, an experimenter pulls the toy out of the green box and shows the children that the bunny can be turned into a carrot-toy by a simple manipulation, before the experimenter puts the toy back into the green box under its carrot aspect. When the agent is back, the children see the experimenter pull the toy under its carrot aspect out of the green box and place it into the other yellow box. Children are directly asked where the agent will look for the bunny. If (and only if) children understand that the agent fails to know that the carrot-toy is the bunny, then they are likely to ascribe to the agent the false belief that the bunny is still in the green box and to correctly predict that she is likely to look for the bunny in the green box.[14]

Early studies by Apperly and Robinson (1998; 2003) found that most children, who succeed on non-aspectual verbal false-belief tasks about an object's location, fail on these aspectual verbal false-belief tasks about object-identity. However, more recent findings by Rakoczy et al. (2015) show that most 4-year-old children, who succeed on verbal false-belief tasks about an object's location, also succeed on such aspectual verbal false-belief tasks about object-identity.[15]

Burge (2018, p. 426) discusses these aspectual studies based on false-belief tasks about object-identity and argues that they fail to provide decisive evidence that children can ascribe false beliefs about object-identity and thereby understand the aspectuality of beliefs, because these studies can be interpreted in terms of his non-mentalistic action-sensing attribution scheme. Burge's (2018, p. 426) basic point is that the mentalistic interpretation of these findings "conflates children's tracking properties with their tracking mental modes of presentation (or representations)." In line with his non-mentalistic account of Song and Baillargeon's (2008) study with the skunk and the doll with blue hair, he surmises that it is possible to account for these findings by assuming that young children attribute generic non-mental sensing to an agent, where "generic sensing can track objects by tracking relevant properties" (Burge 2018, p.

[13] Unlike false-belief tests about an object's location, false-belief tests about object-identity have recently been called "aspectual" studies (cf. Rakoczy et al. 2015).

[14] Cf. Rakoczy, Fizke, Bergfeld, & Schwarz (2015) and Rakoczy (2016) for extended discussion. Burge himself discusses a study involving a toy figure first called Peter and casually dressed, then introduced as "the fire-fighter" and wearing a fire-fighter uniform that makes him look unrecognizably different from when he is casually dressed.

[15] Findings based on the false-belief condition of Scott and Baillargeon's (2009) penguin study (discussed in section 3) can be taken to show that 18-month-olds attribute to the agent the false belief that the assembled two-piece penguin in the transparent box is the one-piece penguin. Buttelmann et al. (2015) provide evidence that most 18-month-olds give a rock, not a sponge, to an agent who is unsuccessfully trying to reach for a sponge that deceptively looks like a rock and which the agent falsely believes to be a rock.

420). In true-belief conditions, children take the agent to track (sense or register information about) a pair of properties of the relevant object (e.g. the properties of being a bunny and a carrot-toy). In false-belief conditions, they take the agent to fail to register one of its pair of properties (e.g. being a carrot-toy).

As I will now argue, application of Burge's non-mentalistic scheme to these aspectual studies has the unexpected effect of making true-belief conditions cognitively more demanding than false-belief conditions. In the false-belief condition, children are said to take the agent to track the toy with two aspects or properties (e.g. a bunny and a carrot aspect) by sensing only one of its aspects or properties (e.g. the bunny aspect). In the true-belief condition, children are said to take the agent to track the very same toy by sensing both of its aspects or properties. This seems to make the true-belief condition more demanding than the false-belief condition. From a mentalistic construal of these findings, it seems counter-intuitive that a condition in which the children and the agent share the same information about the toy would be more demanding than a condition in which children must, as Burge puts it, decouple the information they take to be available to the agent from their own.

6 The Appeal to Parsimony

In what follows, I will grant Burge the claim that the action-sensing attribution scheme is part of stable scientific biological explanations of actions (as recognized at the end of section 2).[16] However, the relevant psychological question is whether human infants apply this scheme to others' actions and to actions performed by non-human animals, including snails and ticks. Since Burge accepts the mentalistic criticisms of all three non-mentalistic accounts of the infant data other than his own action-sensing attribution scheme, he appeals to a principle of parsimony in favor of his action-sensing attribution scheme against its mentalistic alternative. According to Burge's (2018, p. 211) methodological principle of parsimony, which he takes to be in the spirit of Ockham and Morgan, "when infants and non-human animals are known to attribute a certain property (or relation or kind), an explanation that takes them to attribute a further property that is a subspecies of the first is to be rejected, unless it is supported by evidence that shows that relevant subjects have capacities specific to the subspecies."

Given that sensing is a genus of which perception is a species, perceptual states stand to sensory states as subordinate states to super-ordinate states. (Similarly, desires and intentions stand to non-mental generic conative states as subordinate to superordinate conative states.)[17] Given developmental evidence consistent with the hypothesis that infants (and non-human animals as well) attribute superordinate generic states, Burge's parsimony principle stipulates that the burden of proof is on the competing subordinate mentalistic hypothesis. Additional evidence is required to support the more specific mentalistic hypothesis that infants attribute mental states over the competing generic hypothesis. Parsimony raises a contentious issue in the developmental debate

[16] In other words, I will bracket any doubt as to whether a state of sensory retention is a genuine non-mental counterpart to mental states of believing or remembering.

[17] It is not clear, as I have argued in section 3, whether Burge thinks that the distinction between intention and desire (two mental states) can be mapped onto an equivalent distinction between a pair of generic non-mental internal conative states.

about mindreading in human infancy. It has been a contentious issue in a related debate, namely the debate about mindreading in non-human primates. More generally, it can be shown to be a contentious issue in evolutionary theorizing.[18]

Appeal to parsimony in the debate about mindreading in non-human primates was initially introduced by critics of the mentalistic interpretation who advocated a behavior-reading interpretation (cf. Povinelli and colleagues). Consider studies by Tomasello and colleagues, which they interpreted as evidence that non-human apes understand the mentalistic difference between knowing and ignoring something (e.g. the location of food) based on the difference between seeing it or not (cf. Hare, Call, Agnetta & Tomasello 2000; Hare, Call & Tomasello 2001). According to advocates of behavior-reading, explaining why a subordinate chooses food occluded to a dominant by the subordinate's ability to read the dominant's behavior (e.g. bodily posture and eye- or head-direction) is more parsimonious than explaining it by the subordinate's ability to attribute ignorance to the dominant. What is to be explained is the subordinate's own behavior of selecting occluded food over non-occluded food. It is common ground that the subordinate could not attribute mental states (of seeing and knowing or not) to the dominant (in accordance with the mentalistic interpretation) unless the subordinate could form beliefs about the dominant's behavior. Povinelli and colleagues' further claim was that the subordinate's beliefs about the dominant's behavior is all that is needed to explain the subordinate's own behavior. The subordinate's putative beliefs about what the dominant can see and know or not are superfluous.

Parsimony in the mindreading debate about non-human apes is a contentious issue because, as advocates of the mentalistic interpretation have pointed out (e.g. Tomasello and Call 2006), the behavior-reading hypothesis is likely to require a different special explanation for each behavioral task (and postulate a collection of behavioral rules), whereas the mindreading alternative offers a single unified explanation for many different behavioral tasks.

The point is well spelled out by Sober (2016a, b), who combines two studies by Melis et al. (2006). In the "tunnel" study, chimpanzees can look through a window into a room in the middle of which a human is standing. They also can reach inside the room to get at two food items from either an opaque or a transparent tunnel. Melis et al. (2006) found that they reliably chose the opaque over the transparent tunnel. In the "trapdoor" study, they could steal the food from either a noisy or a quiet trapdoor. Melis et al. (2006) found that they reliably chose the quiet over the noisy trapdoor. What Sober (2016a, b) argues is that on the mindreading hypothesis, but not on the behavior reading hypothesis, the two tasks have a common (mentalistic) structure in two different sensory modalities: chimpanzees can be taken to select the channel for stealing the food that minimizes the likelihood that the human will notice (or know) what they are doing and that therefore maximizes the probability of the success of their own action.

Burge's appeal to parsimony in the interpretation of the infant developmental data is contentious for at least two reasons. First of all, as the debate about mindreading in non-human primates makes clear, there are two different senses of parsimony, one arguably derivable from Morgan's canon and the other arguably derivable from Ockham's razor. According to Morgan's canon, we should refrain from interpreting an action "as the outcome of a higher psychical faculty, if it can be interpreted as the outcome of the

[18] Cf. Sober (1998) for extended discussion of this issue in evolutionary theorizing.

exercise of one which stands lower in the psychological scale" (Morgan 1894, p. 53). On Morgan's canon, we should probably prefer the behavior reading interpretation to the mindreading interpretation of the findings about non-human primates, on the grounds that mindreading is a "higher psychical faculty" than behavior reading. However, in virtue of Ockham's razor, we might prefer the mindreading explanation to the behavior reading explanation, on the grounds that the former is more unified than the latter (which appeals to a collection of behavioral rules). Although Burge (2018) never expresses sympathy towards the behavior reading interpretation of the infant data, it is not entirely clear which position in the debate about mindreading in non-human primates he supports. Nor is it clear that he fully recognizes that Morgan's canon and Ockham's razor might support different notions of parsimony.

Secondly, Burge (2018) appeals to parsimony in favor of his own non-mentalistic action-sensing attribution scheme against the mentalistic alternative, not against the behavior reading alternative, in the debate about mindreading primarily in human infancy, not in non-human primates. Presumably, both Burge's non-mentalistic action-sensing attribution scheme and its mentalistic alternative are equally unified attribution schemes. If so, then Ockham's razor provides no support to either alternative in the dispute about how to describe the infant data. If so, then what justifies Burge's (2018) claim that it is more parsimonious to describe the infant data by crediting infants with his action-sensing attribution scheme rather than with the capacity to attribute mental states is some version of Morgan's canon because, as Burge does emphasize, to appeal to mental states is to appeal to more specific (and therefore to a "higher psychical faculty") than to appeal to generic non-mental internal states.

However, there is an obvious asymmetry recognized by Burge (2018, p. 412) between human infants and non-human primates. First of all, we know that human children eventually grow into older children and adults who are able to attribute mental states; but we do not know in the same way that non-human animals do so. So by Burge's own admission, I think, the relevant version of the parsimony principle should not apply merely to a pair of competing unified hypotheses, both of which purport to describe the infant data. It should apply to competing pairs of hypotheses, one of which offers a unified description of the infant data, and the other of which purports to explain how infants become able to attribute mental states to others. On the mentalistic hypothesis, infants can attribute mental states to others. So the further burden of explaining how children are able to attribute mental states to others is now shifted to biological evolution. But so long as Burge does not propose a mechanism explaining how infants acquire the capacity to attribute mental states to others, it is difficult to evaluate his claim that his action-sensing attributing scheme is more parsimonious than its mentalistic alternative. Moreover, on Burge's approach, the burden of explaining how infants are able to apply his non-mentalistic action-sensing scheme must also be shifted to biological evolution.

Secondly, as Burge (2018, p. 412) himself recognizes, human adults (and older children as well) are not just known to mentalize, but to over-mentalize, i.e. to attribute mental states to entities that lack mental states. Not only is there cross-cultural evidence of over-mentalization in various human religious systems around the world, but there is much psychological evidence that human adults are prone to spontaneously over-attribute mental states to e.g. puppets, smurfs or geometrical stimuli that they know to lack mental states (cf. Heider and Simmel 1944; Kovács et al. 2010). On the

mentalistic interpretation (but not on Burge's non-mentalistic attribution scheme), infants too have been shown to be prone to spontaneously over-attribute mental states to puppets, smurfs or geometrical stimuli, whether or not infants know that puppets, smurfs and geometrical stimuli lack mental states. Human adults are also clearly able to reflectively suspend their spontaneous attribution of mental states to entities which they know to lack mental states (e.g. puppets). In fact, it seems likely that human biologists could not come to attribute to snails, ticks and bacteria non-mental generic sensory and conative states unless they could inhibit their spontaneous propensity to attribute mental states to entities which they know to lack mental states could human biologists. It therefore seems odd that human infants would have the spontaneous capacity to attribute to humans generic sensory and conative states and entirely lack the capacity to attribute mental states to others.

7 Concluding Remarks

Burge's action-sensing scheme is a new, interestingly unified, non-mentalistic interpretation of developmental findings that have been interpreted as evidence of mental state attribution in early human childhood. It is an upshot of his important work on the objectivity of perceptual representations. According to Burge's framework, young children fail to discriminate non-intentional actions performed by ticks or snails from intentional actions performed by humans: they interpret all actions in terms of the kind of generic non-mental internal states that biologists selectively attribute to ticks and snails. Burge's (2018) interpretation of the developmental findings based on his non-mentalistic attribution scheme is fully consistent with his strong claim that infants can form objective sensory perceptual representations but lack thoughts and beliefs with propositional contents (Burge 2010). If infants cannot form thoughts and beliefs with propositional contents, then a fortiori they cannot form beliefs about others' thoughts and beliefs with propositional contents.

I have raised two basic kinds of challenges for Burge's non-mentalistic action-sensing attribution scheme. In section 3, I have highlighted particular developmental studies that provide evidence that infants are sensitive to subtle differences among the epistemic mental states of other people: in particular between another's ignorance and another's false belief. On this basis, I have challenged Burge's claim that his non-mentalistic scheme has the resources to match such subtle distinctions among epistemic mental states. In section 4, I have challenged Burge's account of findings based on false-belief tasks about object-identity that many psychologists take to have the potential to demonstrate that children understand the aspectuality of beliefs.

In section 5, I have questioned Burge's claim that his non-mentalistic account of the infant data is more parsimonious than the mentalistic alternative along two different lines. First of all, I have argued that parsimony inspired by Morgan's canon, but not by Ockham' razor, might justify the preference for an account of the infant data in terms of infants' capacity to attribute more generic internal states rather than more specific mental states. Secondly, I have further argued that parsimony considerations should not apply merely to competing descriptions of the infant data. Instead they should apply to competing descriptions of the infant data coupled with hypotheses about the cognitive mechanisms that enable human infants to acquire the capacity to mentalize.

Burge acknowledges that adults over-mentalize, but he misses, I think, the potential significance of this fact for the dispute about the interpretation of the infant data. Human adults are known to spontaneously (if not automatically) attribute mental states, not generic internal states, to entities that they know lack mental states. If so, then they might have to inhibit their spontaneous (or automatic) propensity to mentalize in order to attribute generic internal states to ticks, snails or bacteria. On the mentalistic interpretation of the developmental findings, adults' propensity to over-mentalize can be construed as the adult "signature" of infants' capacity to mentalize. But Burge's non-mentalistic interpretation of the developmental findings seems to turn things upside down: it gives ontogenetic priority to the application of the non-mentalistic action-sensing attribution scheme to human actions: so it takes mentalization to be effortful in early human childhood and even perhaps to require inhibition of the non-mentalistic action-sensing attribution scheme to human actions.

Acknowledgements I am grateful to the Editor and to four anonymous reviewers, to Andreas Falck, George Galfalvi, Gyuri Gergely and Dan Sperber, for their insightful comments and criticisms on earlier versions of this paper, and also to Ned Block for conversations on the topic of this paper. I gratefully acknowledge support from the European Research Council (ERC) under the European Union Seventh Programme (FP/2007-2013)/ ERC Grant 609819.

References

Apperly, I. & Robinson, E. 1998. Children's mental representation of referential relations. *Cognition* 67, 287–309.

Apperly, I. & Robinson, E. 2003. When can children handle referential opacity? Evidence for systematic variation in 5- and 6-year-old children's reasoning about beliefs and belief reports. *Journal of Experimental Child Psychology* 85, 297–311.

Apperly, I., and S. Butterfill. 2009. Do humans have two systems to track beliefs and belieflike states? *Psychological Review* 116 (4): 953–970.

Baillargeon, R., R. Scott, and Z. He. 2010. False-belief understanding in infants. *Trends in Cognitive Sciences* 14: 110–118.

Bloom, P., and T. German. 2000. Two reasons to abandon the false belief task as a test of theory of mind. *Cognition* 77: B25–B31.

Burge, T. 1979. Individualism and the mental. In *Midwest studies in philosophy*, ed. P. French, Theodore E. Uehling Jr., and H.K. Wettstein, vol. IV. Minneapolis: University of Minnesota Press.

Burge, T. 2010. *Origins of objectivity*. Oxford: Oxford University Press.

Burge, T. 2014. Perception: Where mind begins. *Philosophy* 89: 385–403.

Burge, T. 2018. Do infants attribute mental states? *Psychological Review* 125 (3): 409–434.

Buttelmann, D., M. Carpenter, and M. Tomasello. 2009. Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition* 112: 337–442.

Buttelmann, F., J. Suhrke, and D. Buttelmann. 2015. What you get is what you believe: Eighteen-month-olds demonstrate belief understanding in an unexpected-identity task. *Journal of Experimental Child Psychology* 131: 94–103.

Butterfill, S., and I. Apperly. 2013. How to construct a minimal theory of mind. *Mind & Language* 28: 606–637.

Carey, S. 2009. *The origin of concepts*. Oxford: Oxford University Press.

Carruthers, P. 2013. Mindreading in infancy. *Mind & Language* 28: 141–172.

Carruthers, P. (manuscript) Representing the mental as such in infancy.

Davidson, D. 1982. Rational animals. In *Davidson, D. (2001) subjective, intersubjective, objective: Philosophical essays*, vol. 3. Oxford: Oxford University Press.

Dretske, F. 1988. *Explaining behavior*. Cambridge: MIT Press.

Dretske, F. 1995. *Naturalizing the mind*. Cambridge: MIT Press.

Fabricius, W.V., T.W. Boyer, A.A. Weimer, and K. Carroll. 2010. True or false: Do 5-year-olds understand beliefs? *Developmental Psychology* 46 (6): 1402–1416.

Frege, G. 1892/1952. On sense and reference. In *Translations from the Philosophical Writings of Frege by Max Black and Peter Geach*. Oxford: Blackwell.

Gergely, G., and G. Csibra. 2003. Teleological reasoning in infancy: The naive theory of rational action. *Trends in Cognitive Sciences* 7: 287–292.

Hare, B., Call, J., Agnetta, B., & Tomasello, M. 2000. Chimpanzees know what conspecifics do and do not see. *Animal Behaviour* 59(4), 771–785.

Hare, B., Call, J., & Tomasello, M. 2001. Do chimpanzees know what conspecifics know? *Animal Behaviour* 61(1), 139–151.

He, Z., M. Bolz, and R. Baillargeon. 2011. False-belief understanding in 2.5-year-olds: Evidence from violation-of-expectation change-of-location and unexpected-contents tasks. *Developmental Science* 14 (2): 292–305.

Heider, F., and M. Simmel. 1944. An experimental study of apparent behavior. *American Journal of Psychology* 57: 243–259.

Heyes, C. 2014a. False belief in infancy: A fresh look. *Developmental Science* 17: 647–659.

Heyes, C. 2014b. Submentalizing: I am not really reading your mind. *Perspectives on Psychological Science* 9: 131–143.

Hyde, D.C., M.A. Aparicio Betancourt, and C.E. Simon. 2015. Human temporal-parietal junction spontaneously tracks others' beliefs: A functional near-infrared spectroscopy study. *Human Brain Mapping* 36: 4831–4846.

Hyde, D.C., C.E. Simon, F. Ting, and J. Nikolaeva. 2018. Functional organization of the temporal-parietal junction for theory of mind in preverbal infants: A near-infrared spectroscopy study. *Journal of Neuroscience* 38 (18): 4264–4274.

Knudsen, B., and U. Liszkowski. 2012. 18-month-olds predict specific action mistakes through attribution of false belief, not ignorance, and intervene accordingly. *Infancy* 17 (6): 672–691.

Kovács, A., E. Téglás, and A. Endress. 2010. The social sense: Susceptibility to others' beliefs in human infants and adults. *Science* 330: 1830–1834.

Leslie, A. 1994. ToMM, ToBY and agency: Core architecture and domain specificity. In *Mapping the mind: Domain specificity in cognition and culture*, ed. L. Hirschfeld and S. Gelman. New York: Cambridge University Press.

Luo, Y. 2010. Do 8-month-old infants consider situational constraints when interpreting others' gaze as goal-directed action? *Infancy* 15 (4): 392–419.

Luo, Y. 2011. Three-month-old infants attribute goals to a non-human agent. *Developmental Science* 14: 453–460.

Luo, Y., and R. Baillargeon. 2005. Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science* 16: 601–608.

Melis, A., J. Call, and M. Tomasello. 2006. Chimpanzees (Pan troglodytes) conceal visual and auditory information from others. *Journal of Comparative Psychology* 120: 154–162.

Millikan, R.G. 2004. *The varieties of meaning*. Cambridge: MIT Press.

Morgan, C. Lloyd. 1894. *An introduction to comparative psychology*. London: Walter Scott.

Oktay-Gür, N. & Rakoczy, H. 2017. Children's difficulty with true belief tasks: Competence deficit or performance problem? *Cognition* 166, 28-41.

Onishi, K., and R. Baillargeon. 2005. Do 15-month-old infants understand false beliefs? *Science* 308: 255–258.

Penn, D., and D. Povinelli. 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* 362: 731–744.

Perner, J., and J. Roessler. 2012. From infants' to children's appreciation of belief. *Trends in Cognitive Sciences* 16 (10): 519–525.

Perner, J., and T. Ruffman. 2005. Infants' insight into the mind: How deep? *Science* 308: 314–316.

Povinelli, D., and T. Eddy. 1996. What chimpanzees know about seeing. *Monographs of the Society for Research in Child Development* 61 (3): 1–152.

Povinelli, D.J., and J. Vonk. 2004. We don't need a microscope to explore the Chimpanzee's mind. *Mind and Language* 19: 1–28.

Rakoczy, H., D. Bergfeld, I. Schwartz, and E. Fizke. 2015. Explicit theory of mind is even more unified than previously assumed: Belief ascription and understanding aspectuality emerge together in development. *Child Development* 86 (2): 485–502.

Scott, R., and R. Baillargeon. 2009. Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development* 80: 1172–1196.

Sober, E. 1998. Morgan's canon. In *The evolution of mind*, ed. D. Dellarosa Cummins and C. Allen, 224–242. Oxford: Oxford University Press.

Sober, E. 2016a. *Ockham's razors: A user's manual*. Cambridge: Cambridge University Press.

Sober, E. 2016b. Replies to Kristin Andrews's, Gordon Belot's, and Patrick Forber's reviews. *Metascience* 25: 393–403.

Song, H., and R. Baillargeon. 2008. Infants' reasoning about others' false perceptions. *Developmental Psychology* 44: 1789–1795.

Spelke, E. 1988. Where perceiving ends and thinking begins: The apprehension of objects in infancy. In *Perceptual development in infancy, the Minnesota Symposia on child psychology*, ed. A. Yonas, vol. 20. Hillsdale: Lawrence Erlbaum.

Tomasello, M., and J. Call. 2006. Do chimpanzees know what others see—or only what they are looking at? In *Rational animals*, ed. S. Hurley and M. Nudds. Oxford: Oxford University Press.

Wellman, H.M., Cross, D. and Watson, J. 2001. Meta-analysis of theory of mind devel- opment: the truth about false belief. *Child Development* 72, 655–84.

Wimmer, H., and J. Perner. 1983. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13: 103–128.

Woodward, A. 1998. Infants selectively code the goal object of an actor's reach. *Cognition* 69: 1–34.