Representing the mental as such in infancy

Peter Carruthers
University of Maryland

Abstract: This paper critiques the claim made by Burge (2018), that none of the existing evidence for mental-state attribution prior to the age of 4 or 5 really does support such a conclusion (not just for beliefs, but for mental states of all sorts). In its place, he offers an explanatory framework that attributes mere information-states and teleologically-characterized goals, which are said to lack the defining properties of the mental. I argue that Burge’s claims are poorly motivated and misrepresent the goals of developmental psychology. I offer in their place an interpretation according to which infants and young children increasingly represent the mind as such, but do so by degrees.

Tyler Burge (2018)\(^1\) mounts a systematic attack on the widely accepted claim that infants, young children, and nonhuman animals represent the minds of others, or at least some aspects of the minds of others. His argument is that while members of these groups may be representing what are, in fact, mental states in other agents, those states aren’t represented as mental ones. That is, infants and young children don’t represent mental states as such.

Burge stipulates at the outset (without evidence) that it is agreed among philosophers that in order to qualify as mental, a state must either be conscious or representational, or both. The intended status of this claim is unclear, however, at least initially. Is the idea that the disjunction is a necessary condition on mentality? Or is either one of the two disjuncts thought to be individually sufficient as well as disjointly necessary? Throughout most of Burge’s discussion the answer appears to be the latter. For he focusses almost entirely on the representation-clause, and appears to treat someone representing a state as representational as the “gold standard” for representing minds. Certainly he thinks that without representing a state as representational one isn’t representing the mental as such (unless one is representing a conscious sensation like pain, which Burge thinks isn’t really representational). And for much of the discussion it seems to be assumed that passing this test would be sufficient for a being to represent mindedness, also.

It emerges quite late in Burge’s discussion, however, that representation is only a necessary

\(^1\) All references to Burge are to this paper unless otherwise noted.
condition on mentality (or rather, is one part of a disjunctive necessary condition). Representational states need to be embedded in the right kind of causal network to qualify as mental, Burge tells us (p.425). This is surely correct. One can imagine a camera-like mechanism, for example, that is designed just to extract visual color constancies from the vagaries of lighting conditions (this is Burge’s paradigm of a representational state), but without having the capacity to do anything further with them. Although qualifying as representational by Burge’s lights, this plainly wouldn’t qualify the states of the camera as mental.

There is another ambiguity, or apparent ambiguity, in Burge’s presentation from the outset. Are the conditions on mentality that he puts forward conditions on real minds, as scientifically understood? Or are they conditions on mentality as ordinary folk understand the mental? For his purposes, he plainly needs to intend the latter. For it would be no news to anyone that young children fail to represent mental states in the way that cognitive scientists do, any more than it is news to claim that children’s understanding of physics differs from that of physicists (Carey 2009). Burge’s claim is about the course of normal human development, and his main question should thus concern the stage at which children first represent mental states as mental in whatever way the latter notion is understood by ordinary unscientific adults.

In fact, however, Burge elucidates the notion of representation, and the need for explanatory appeals to representation, within a cognitive-scientific framework. He tells us at the outset that representations need to be capable of being accurate or true, and that appeals to correctness conditions, in turn, need to figure in the law-like explanations provided by a stable science of the mental. He points out that the informational states of a slug, for example, on detecting a nearby lettuce leaf, aren’t regarded by cognitive scientists as genuinely representational, because such a description would do no explanatory work. Rather, we can say that the slug has a primitive drive to move toward the leaf, activated by sensory/informational triggers.

Admittedly, Burge does also say that the information-based explanatory scheme that he suggests young children and some animals employ to predict and explain one another’s behavior is also available to common sense. But he offers no evidence that it is actually employed by the folk. And as the discussion and development of the non-mentalistic scheme is developed and elaborated in the course of Burge’s paper, it proves to be quite hard to deploy, except laboriously and explicitly. Although it may be possible to induce ordinary folk to employ simple versions of it (for example, to explain the behavior of the slug), this suggests that the full scheme is not within their normal repertoire. Indeed, even coming to recognize the possibility of a non-mentalistic internal-state scheme of explanation and prediction seems
to be an effortful intellectual achievement. (At any rate, it was so for the present reader.) This may well be why nearly everyone in the field moves straight from rejection of a mere behavior-rule account of infant competence to a mentalizing account. This would appear to suggest that it is mental-state attribution that is cognitively primitive, and has to be scaled back by effortful scientific inquiry so that one doesn’t attribute mentality to slugs.

Burge may well be correct that slugs don’t genuinely represent the world around them, and don’t qualify as having mental states (at least, in any scientifically realistic sense). But the pertinent question is: what do the folk think? For what should be at stake, remember, is the notion of the mental that regular (unscientific) adults employ. The goal is to understand when children achieve an adult-like competence (not when, if ever, they achieve a scientific one). Burge provides no evidence that the folk would deny mentality to the slug. And on the contrary, attributing mental states to it feels quite natural, even when one is fully informed of the limitations on the slug’s behavior. One might say, “The slug smells the leaf nearby, and moves toward it because it wants to eat it.” But if the slug has mental states by the lights of the folk, then Burge’s argument collapses. For his minimal (non-mentalistic) informational scheme will actually be mental by the lights of the folk. And then by those same lights (if we accept the infancy data as valid), human infants represent others’ mental states as mental from at least early in the second year of life.

A similar difficulty infects Burge’s examples of merely-informational (non-representational) sensing. He says, for example, that bees use non-representational sensing of visual flow to determine distance of flight (p.410, fn.4). And registration of the light array striking the retina is said not to genuinely represent the ambient patterns of light (p.410). Both of these claims are said by Burge to be scientifically correct. And perhaps he is right, from a scientific perspective. But my guess is that the folk find it quite natural to think otherwise, and would maintain that the states in question possess correctness-conditions. For instance, on learning that experimenters can lead bees to underestimate the distance travelled by varying the level of contrast in the patterns on the walls of the tunnel they fly down (as reflected in their subsequent waggle dance; see Si et al. 2003), one might say, “They misrepresented the amount of visual flow, and that is why they got the distance wrong.” And someone with red-green color blindness might naturally say, “I often get the colors of things wrong because my

---

2 Exceptions include those who defend dual-system theories of action-explanation (Apperly & Butterfill 2009), who postulate that the explanatory scheme employed by infants and young children doesn’t really represent mental states as such, but continues to exist alongside theory-of-mind abilities in adulthood. Burge makes no such claim.
retina doesn’t tell me about the difference between red and green”. (Note that instead of “doesn’t tell me about”, one could equally say, “doesn’t represent”. ) It seems likely, then, that the folk notion of representation can apply to instances of what Burge calls “mere sensing”. And in that case the states in question satisfy at least one component of the folk notion of mentality.

The question whether, and when, children acquire a scientific understanding of the mental is a legitimate target of inquiry, of course. Similar questions have been addressed in other domains of knowledge, especially physics and biology where children’s naïve theories can interfere with their grasp of important aspects of the mandatory school curriculum (Clement 1982; McCloskey 1983; Shtulman & Valcarcel 2012). But psychology and cognitive neuroscience aren’t generally taught in school, so the corresponding question has not seemed urgent. And certainly it hasn’t been the goal of the extensive theory-of-mind literature. On the contrary, what developmental psychologists have been seeking to understand is how and when children acquire something resembling an adult (common-sense) understanding of the mind.

I have argued that it is problematic for Burge to impose a cognitive-scientific notion of representation as a constraint on children’s capacity of represent mental states as such. It is a mistake to insist that children need to be representing agents as entertaining representations in whatever sense cognitive science construes the latter notion. But there is actually a prior methodological difficulty with the structure of Burge’s argument. For he appeals at the outset to an (alleged) consensus among philosophers to ground his disjunctive condition on mentality. He tells us that philosophers are entirely agreed that to qualify as mental, a state has to be either conscious or representational or both. But the problem here is that the intense philosophical focus on both phenomenal consciousness and mental representation over the last 40 years has occurred as part of a very different kind of project, not as part of an attempt to characterize the folk conception of the mental as such.

On the contrary, the problem in each case has been to understand how the properties in question form part of the natural order of the world, which is taken to be ultimately physical in nature. In connection with consciousness, there have been extensive debates about the so-called “hard” problem and the explanatory gap (Dennett 1991; Tye 1995, 2000; Chalmers 1996; Carruthers 2000; Strawson 2006). And in connection with representation, likewise, philosophers have tried to understand how representation can be a natural (ultimately physical) property of physical agents (Millikan 1984; Papineau 1987; Dretske 1988; Fodor 1990). Moreover, many of those who have addressed the latter question have been especially focused on the way representation should be characterized for the purposes of cognitive science itself. Part of their interest has been in the question whether some
appropriate notion of representation will ultimately form part of a mature science, or will rather be eliminated from science.

These discussions have definitely not taken place in the context of trying to characterize necessary and/or sufficient conditions on mental states as such (and hence on what it would take to represent a state as mental). Even less have they formed part of an attempt to characterize the folk-notion of the mental. Rather, those discussions have focused on these two properties (phenomenal consciousness and representational content) as especially puzzling properties that some mental states possess.

In order to find philosophical discussions of the nature of the mental as such, one has to go back to an older literature, on functionalism and naïve theories (Lewis 1966, 1970; Putnam 1967, Armstrong 1968). Here something like a consensus did emerge (at least, if the problem of consciousness is set to one side, as for the most part Burge himself does). It is that what constitutes a state as a mental one is its place in the causal system described by our folk theory of the mind, where this can be thought of as constituted by, or embedded in, the set of platitudes (whether explicit or merely tacit) that ordinary people accept about mental states, their causal roles, and their other properties.3

Moreover, psychologists who investigate the development of the folk understanding of the mind have almost universally converged on some version of “theory-theory” (Wellman 1990; Perner 1991; Gopnik & Wellman 1992; Wellman et al. 2001; Baillargeon et al. 2010). They have differed over whether, or how much of, the folk theory of the mental is explicitly represented or left implicit in the processing principles employed. And they have differed over how much of the theory is innately specified, as well as over the manner in which normal development proceeds from whatever is the initial starting-state. But all agree that the upshot is a theory-like structure. Even those who have defended introspection-based simulation, rather than theorizing, as the main engine of acquisition acknowledge that the outcome of development includes a body of theory-embedded concepts and causal principles (Goldman 2006).

3 Seen in this light, it isn’t clear that phenomenal consciousness would figure in the folk-definition of mentality at all. For it is a term of philosophical art. The distinction between phenomenal (or “feely”) mental states and others is pretty much invisible from the perspective of common-sense psychology. Agents can be conscious as opposed to asleep, of course; and likewise the folk make use of the notion of an agent being conscious of some thing or event, meaning that the agent perceives it. But neither of these is the same as phenomenal consciousness, which is thought by philosophers to be an introspectively-accessible—and puzzling—property that some mental states possess.
According to the folk-theory account of the mental, then, the representational properties of (some) mental states will be one sort of property of the mental among others. It is unclear that they should be accorded any special status. Indeed, like theories generally (whether scientific ones, or the common-sense theories of the folk) it is doubtful that any of the properties described by the theory gets picked out as strictly necessary for the theory to apply. Rather, while some properties and some causal principles may be more central to the theory than others, all have essentially the same causal-law-like status. And systems that possess only some of these properties, or satisfy some of the causal generalizations, will then qualify as instances of the completed theory to some significant degree. Given the likely structure of folk-psychology, then, having a mind, and being a mental state, will be matters of degree. And we would expect that infants and young children acquiring the theory piecemeal over time could be said to be representing the mental states of other agents as mental to some significant degree or other.

As already remarked, Burge’s focus is, almost exclusively, on children’s understanding of the representational character of the mental (and the lack of evidence thereof). He doesn’t tell us what it takes for a mental state to qualify as a representation, however. He has no theory of representation. Indeed, we know from his 2010 book that he believes it is a mistake to seek a reductive account of representation. This is a view I am quite sympathetic towards (Botterill & Carruthers 1999). The notions of representation and intentional content are validated by their foundational role in cognitive science, and we should look to an analysis of those sciences to see where those notions are genuinely needed. This is why Burge (2010) focuses especially on visual constancies: they are, he argues, the earliest stage in visual processing where the idea of correctness-conditions does genuine scientific work. But now, in the present context, we can see how this might have led Burge (2018) to go astray. He should have been asking at what stage children start working with whatever notion of representation regular folk employ; instead, he seems to be asking when they first employ the notion that cognitive scientists use. Perhaps these are the same. But Burge doesn’t demonstrate that they are; nor does he offer any evidence that they are.

Although Burge (2010, 2018) eschews a theory of representation, he plainly thinks that the content of a representation is either a proposition with truth-conditions, or a contribution to the content of such a proposition. He also plainly conceives of propositions as fine-grained (so-called Fregean) entities. On this view, the proposition that Clark Kent has left the building is distinct from the proposition that Superman has left the building, even though Clark Kent is Superman, and even though most philosophers think that identities like this are necessarily true (true in all possible worlds). But
again, this conception of propositions is by no means universally accepted, even among philosophers. Many think that propositions are coarse-grained entities, constituted, for example, by sets of possible worlds (the worlds at which they are true). On this account, the two propositions about Clark Kent / Superman are actually one and the same, and what differentiates them is merely pragmatic (Lewis 1986; Stalnaker 1999; Saul 2007). Yet what should matter for Burge’s purposes, of course, isn’t what philosophers think, but rather what the folk do. He provides us with no evidence on that front, either.

Central to Burge’s argument against infant/childhood mentalizing is the example of the slug, which has informational states caused by properties in the world (the odor emitted by a lettuce leaf, for instance) and simple drives that can be triggered by the presence of such states. The slug, Burge thinks, doesn’t have mental states, because it doesn’t truly represent the world around it. The states that infants and young children attribute to other agents need be no different in principle from those attributed to slugs. Although children develop an increasingly sophisticated explanatory scheme, involving informational states, stored information, drives, and teleologically-characterized goal states, this is said be to just an elaboration of the scheme that can be used to explain the behavior of the slug. Critically, this scheme doesn’t employ the notion of representation, and hence is said not to involve attributions of mental states as such.

Suppose we grant that slugs lack minds, however, and grant that the folk would think that slugs lack minds. What would need to be added to the slug’s inner life or behavior for it to qualify as having representations of the world around it? On one quite intuitive approach, it would be for the states in question to be embedded in a more flexible inferential system. On this view, what transforms a mere information-carrying state into a mental state with representational content, is nothing intrinsic to that state as such, but rather its embedding within an intelligent system. Thus, people might be disinclined to attribute mental states to slugs because slugs aren’t smart enough, and aren’t flexible enough in their behavior. But if so, as more and more complexity gets added to the information-registering scheme (of the sort that Burge acknowledges infants and young children can represent in others) then it is far from clear that the states in question don’t turn into genuinely representational ones (from the perspective of common sense) after all.

In fact, one might think that the interlocking structure of belief and desire is just as central to folk psychology as is the notion of representation, and this is arguably represented from the earliest stages of infancy. This scheme will include principles such as that someone who wants something will do what he thinks will get him that thing. The idea of potential misrepresentation may be tacit here. For it includes at least an implicit appreciation that if the agent is wrong about what will achieve his goal, then
the action will fail, and the desire will go unsatisfied. But the notion of correctness-conditions may not at this stage play a foundational explanatory role.

Notice that this gradualist picture is exactly what one might predict from the perspective of the functionalist / naïve-theory account of the mental. As more and more folk-psychological platitudes and causal generalizations need to get added to explain the behavior of the creature in question, so we should become more and more confident about regarding that creature as possessing a mind. And then by the same token, as infants employ increasingly sophisticated predictive and explanatory schemes, we should count them as representing the states of other agents as mental ones to a greater and greater degree.

Notice, too, that there are a family of theories of representational content (so-called “dual factor” theories) that might generate the same result. On this sort of view, what differentiates between the contents of two representations that carry the same information about a given object or event (such as Clark Kent / Superman) are the ways one’s cognitive system is apt to respond to and make use of those representations. Each may give rise to distinct inferences (“will make nerdy remarks” versus “can leap over buildings”), and will be embedded in different sets of ancillary information or “object files”. So what may transform a given informational state into a fine-grained Fregean representation is its embedding in a cognitive system that is apt to make use of it in flexible fine-grained ways.

Once one thinks of mental states as whatever get described by the overall folk theory of the mind, then one can accept that infants from the earliest stages are representing mental states as mental to some degree. That they are representing states that are, in fact, mental we (and Burge) can agree. But what it takes to represent them as such can be a matter of degree, as infants and young children enrich their principles for attributing states to others and for generating expectations for others’ behavior. Seen in this light, the capacity to represent the states of another agent as representational is just one step among many. And it, too, will likely come in degrees, rather than being categorically present or categorically absent.

Consider a simple change-of-location false-belief task, of the sort that even quite young infants can solve, as measured by looking time (Onishi & Baillargeon 2005; Baillargeon et al. 2010) or anticipatory mirror-neuron activity (Southgate & Vernetti 2014). Whatever representation infants employ to index the previous location of the object to the target agent (whether this is AGENT THINKS, on the one hand, or AGENT HAS STORED INFORMATION, on the other), the fact that the state in question has correctness or truth conditions can be left implicit in the updating principles that the infant employs. In the true-belief condition, for example, the infant first encodes AGENT THINKS: THE BALL IS IN THE BOX, and
then updates that to AGENT THINKS: THE BALL IS IN THE BASKET when the ball is moved in the agent’s presence, forming expectations for the agent’s search-behavior accordingly. In the false-belief condition, the former attribution is left untouched (not updated) when the ball is moved during the agent’s absence. And then that attribution, in combination with AGENT WANTS THE BALL, will lead the infant to expect the agent to move toward the box. At this point in development the infant need not be representing the information-state of the target agent as being the sort of state that has correctness conditions. But still the infant represents it as fitting into an information-desire explanatory scheme, and hence to that degree represents the state as a mental one, on the sort of gradualist account I am proposing.

Now consider an active helping task of the sort conducted by Buttelmann et al. (2009). The infant has watched the agent play with a toy before placing it in box A. In the agent’s absence the toy is switched to box B and both boxes are locked, before the agent returns and attempts to open box A without success. The infant is encouraged to help. Those who succeed move to open box B, reasoning that box B is where the desired toy really is. (In the true-belief condition, in contrast, infants who pass help the agent to open box A, reasoning that he must be wanting to open if for some other reason, since the toy is known to be in B.) In order to succeed in this task, the infant needs to do more than just track and reason from the agent’s belief state or information state. In order to figure out what counts as helping, in these circumstances, the infant has to reason not only that the agent is trying to open box A because that is where he thinks the toy is, but needs to put this together with her own knowledge that the toy is in B, in order to realize that helping to open box B, and not box A, is what is needed for the agent to achieve his goal. So the infant must deploy both the thought: AGENT THINKS: THE TOY IS IN A and her own thought: THE TOY IS IN B together in the same inference. But this just is to represent the agent’s belief as false. (Agent falsely believes that p = not-p & agent believes that p.)

Burge discusses this experiment briefly, and asserts that it can be handled using his generic informational scheme together with the infant’s desire to help the agent achieve his goal. But why would the infant have such a desire, if a goal is conceived of merely as a teleologically-characterized end-state of action? Doesn’t the infant need to appreciate that the agent will be unhappy or sad if he opens the wrong box and fails to obtain the toy? Indeed, it is a weakness of Burge’s discussion generally that the way in which he characterizes representations (in terms of correctness conditions or truth conditions) cannot apply to desires. Desires have satisfaction conditions, not correctness conditions. And we know that children this age already appreciate something of this: they know that people who get what they want will be happy, and that those who fail to get what they want will be sad (Wellman et al., 2000; Taumoepeau & Ruffman 2006). Admittedly, merely knowing this about the satisfaction-
conditions of desire is compatible with a coarse-grained conception of representational-content, rather than Burge’s preferred fine-grained one. Still it seems that a notion of representation is in play at this age, and hence that infants are representing the states in question as mental to some additional degree.

Surprisingly, Burge has little to say about the various misleading-appearance tasks that have been conducted with infants and young children. For one might think that these involve just the sorts of fine-grained correctness-conditions that he regards as definitive of the mental. He does briefly mention the “which Penguin?” study of Scott & Baillargeon (2009), saying only that it can be solved through an attribution of non-mentalistic sensing of tracked objects (p.420). It is far from clear that this claim is correct, however, since the test-conditions are created in the agent’s absence, in which case no tracking can take place. Rather, the task seems to require infants to reason about what the agent will expect given previous exposure to the properties of both a divisible and indivisible penguin.

Burge then discusses Song & Baillargeon (2008) at slightly greater length (pp.420-421). In this experiment a doll with blue hair that the agent has been playing with is placed in one box in the agent’s absence, while the lid of the other has blue hair sticking out misleadingly from beneath it. Infants expected the agent to mistakenly go to the latter box to retrieve the doll. Burge says that this just requires infants to attribute informational sensing of “chained” properties, expecting the agent to approach the blue-hair box not because she thinks the doll is in there, but because she has stored an informational link between the doll and blue hair. Although the explanation may be a possible one, notice how hard it is to think through on one’s own behalf (from the perspective of common-sense psychology) how the explanation is supposed to go. A mentalizing construal is much more natural.

Burge doesn’t discuss the misleading-appearance tasks conducted with eighteen-month-old infants by Buttelmann et al. (2014, 2015), which are arguably more challenging for him to account for. In the false-belief condition of Buttelmann (2015), for example, the experimenter is attempting (but failing) to reach an object on the shelf that looks like a toy bath-duck but is really, as the infant knows, a scrubbing brush. When encouraged to help the experimenter by giving her one of two alternative objects in place of the out-of-reach one—one of which is a toy bath-duck resembling the one on the shelf, and the other of which is a very different looking scrubbing brush—most infants choose the duck. Yet in the true-belief condition, in contrast, most choose the scrubbing brush. In order to figure out that the experimenter wants a bath-toy in the false-belief condition, it seems the infant must realize that she is representing the brush as a toy bath-duck (because that is what it looks like); whereas in the true-belief condition the infant needs to realize that the experimenter is representing the object as a scrubbing-brush, although it looks like a toy duck.
Overall, then, if we are clear that the target of development is the theory-like network of mental-state concepts and beliefs possessed by normal adults, rather than Burge’s stipulative definition, what we get is a picture of the gradual emergence of infants’ capacities to represent mental states as mental. (And by the same token, other animals can represent some aspects of the mental, or represent mental states as such to some degree.) As infants acquire more mental-state attribution principles, as well as additional procedures for predicting and explaining behavior, and as they gradually come to appreciate the representational properties of mental states (first implicitly, then as coarse-grained, then as fine-grained), so it becomes increasingly true that they are representing mental states as such.

Burge nevertheless can explain the data by attributing to infants and young children his non-mentalistic informational scheme, of course. But it seems the only reason to accept this explanation in place of the gradual-acquisition one, would be if it could be shown that the non-mentalistic explanatory scheme continues to exist alongside the mental-state one in adulthood. This is precisely the grounds on which dual-system theorists defend their own non-mentalistic interpretation of the infancy data (Apperly & Butterfill 2009; Butterfill & Apperly 2013). But there are deep problems with the evidence presented in support of the continued existence of their non-mental scheme, which has been critiqued on both theoretical and experimental grounds (Carruthers 2016, 2017; Elekes et al. 2016; Westra 2017; Király et al. 2018). And in any case Burge himself makes no such claim, and offers no such evidence.

Burge’s central claim is that it is an over-interpretation of the data to suppose that young children attribute mental states to other agents as such. But this claim just embeds one of the main assumptions I have challenged in this article. For it assumes a categorical all-or-nothing conception of the mental. If we suppose, in contrast (and as I have argued here), that the common-sense notion of the mental is theory-like, then satisfying the postulates of that theory can be a matter of degree, and children’s increasingly firm grasp of the theory can also be a matter of degree. Indeed, asking as Burge does, “When do children first become capable of representing mental states as such?” can be seen to be a mistake. What we should be asking (and what, in my view, the field of developmental psychology actually is asking) is at what stage infants and young children represent mental states as such to some or other degree, and how and when they increasingly approach a mature folk-understanding of the mental.

References


