

# Commentary on “How to Construct a Minimal Theory of Mind”

Shannon Spaulding  
Assistant Professor, Philosophy  
Oklahoma State University

## 1. Introduction

In the target paper, Stephen Butterfill and Ian Apperly (B&A, for short) urge the need for a theoretical middle ground between full-blown theory of mind cognition and social cognition based on non-mentalistic behavioral rules. B&A argue that we need an account of a *minimal theory of mind*, which could explain social cognitive abilities of human adults under cognitive load, human infants, and some non-human animals. After explaining the motivation for such an alternative, B&A articulate their own account of minimal theory of mind.

B&A surely are right that we need a middle-ground account between full-blown theory of mind cognition and mere behavioral rules. Several theorists have attempted to construct such accounts, including Apperly and Butterfill who, in a previous paper that developed a two-systems account for tracking beliefs and belief-like states.<sup>1</sup> B&A’s current account is an improved version of the previous account. This version is not explicitly committed to two (and only two) systems for theory of mind. The current account consists in 5 (or perhaps 4) principles that go beyond mere behavioral rules but are not constitutive of full-blown theory of mind. This minimal theory of mind account is a fruitful and welcome addition to the literature. B&A correctly identify a theoretical gap and offer an account to fill that gap. My commentary here is meant to help improve the account.

In the next section, I will summarize the concepts and principles of B&A’s minimal theory of mind. Though these concepts and principles are explicit and clear in the article, it will be helpful to have them all laid out in one place. In section 3, I will consider whether minimal theory of mind is a genuine alternative to behavioral rules. In the final section, I will address a minor objection.

## 2. Terminology

The following terminology and principles are crucial for understanding B&A’s account, and I will refer back to them in section 3 where I critique the account.

**Goal:** The outcome to which a behavior is directed.

---

<sup>1</sup> Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states. *Psychological Review*, 116(4), 953.

**Function:** An action has the function of bringing about a goal, *g*, just in case this action has brought about *g* in the past and this action occurs in part because of this fact.

- **Representing a goal:** Consists in representing an outcome as the function of bodily movement, i.e., linking bodily movements to outcomes.
- This contrasts with representing an *intention as such*.

**Encounter:** A relation between an agent and an object in her field. An individual encounters an object in her field just in case that object has been in certain spatial and temporal proximity to the agent.

- **Representing encountering:** Consists in representing relations between an agent and an object.
- This contrasts with representing *seeing (or perceiving) as such*.

**Registration:** A relation between an individual, object, and location. An individual registers an object at a location just in case that agent recently encountered that object at that location. Correct registration is when the object is in the location where the agent registers it.

- **Representing registration:** Consists in representing an agent's recent encounter of an object. Representing correct or incorrect registration depends on whether the object is where the agent last encountered it.
- This contrasts with representing (*truly and falsely*) *believing as such*.

**Five principles of minimal theory of mind cognition:**

1. Bodily movements form units that are directed to goals.
2. One cannot goal-directly act on an object unless one has encountered it.
  - a. If an outcome involves a particular object and the agent has not encountered that object, then that outcome cannot be a goal of her actions.
3. An individual correctly registers an object at a location if and only if she most recently encountered it at that location and the object is in that location. Correct registration is a condition of successful action.
4. When an agent performs a goal-directed action with a goal that specifies a particular object, the agent will act as if the object were in the location she registers it in.
5. I'm not entirely sure what the fifth principle is meant to be. In section 4.5, B&A discuss extensions and variations on these principles.
  - a. Extending registration to include other types of property as relations.
  - b. Extending this account to include a relational proxy for desire.
  - c. Varying the types of properties represented (e.g., types of food rather than particular objects).

### 3. Is Minimal Theory of Mind a Genuine Alternative to Behavioral Rules?

B&A present their minimal theory of mind as an alternative to full-blown theory of mind cognition and behavioral rules approaches. The contrast with full-blown theory of mind cognition is particularly clear. Their account does not rely directly on ordinary psychological concepts such as intention, seeing, and belief, like full-blown theory of mind. Moreover, B&A insist that these concepts do not involve metarepresentation, which is central to the full-blown theory of mind account. Finally, these concepts are not merely simplified versions of ordinary psychological concepts. Thus, B&A's minimal theory of mind clearly differs substantially from full-blown theory of mind cognition

B&A's account does not differ substantially from the behavioral account, though. The behavioral rules proponent explains a certain subject's social cognitive behaviors by appealing to behavioral rules or principles that guide the subject's interactions. For example, "People look for objects where they last left them." Employing such behavioral rules does not involve attributing mental states to a target, nor does it involve any sort of metarepresentation. It involves merely attributing a relation between an agent, an object or event, and (sometimes) a location. Similarly, B&A insist that employing their principles does not involve metarepresentation. Their principles involve only relations between agents, objects, and locations. It seems that a behavioral rules proponent consistently could adopt all of the principles B&A advance. Taking their definitions of these concepts and principles at face value, there is nothing incompatible with the behavioral rules approach. Reviewing the concepts and principles described above, it is clear that, though behavioral rules proponents may not currently advance this sort of account, they certainly could take it as a friendly amendment to their account. If this is right, then B&A's account is not a genuine competitor to the behavioral rules account.

If B&A regard their account as a genuine competitor both to full-blown theory of mind cognition and behavioral rules, one thing B&A could do is to explain what feature of their account is incompatible with the behavioral rules approach. Alternatively, they could amend their account so that it is metarepresentational, but not in the way that the robust theory of mind account is. They could regard representing goals as representing a representation, yet regard this as distinct from representing *intentions*.

B&A distinguish between representing a *goal* and representing an *intention*.<sup>2</sup> They also distinguish between representing a mental state and representing that mental state *as such*. Combining these two distinctions, we get the following 4 options:

---

<sup>2</sup> They also distinguish encountering/perceiving and registering/believing. For simplicity, I shall use just the goal/intention distinction.

1. Representing a goal
2. Representing a goal *as such*
3. Representing an intention
4. Representing an intention *as such*

It is not always clear how to make sense of these 4 options with respect to B&A's distinction between theory of mind *cognition* and theory of mind *ability* (pp. 2-3). B&A hold that full-blown theory of mind cognition involves representing propositional attitudes as such (option 4). But, beyond this the distinction is unclear. Presumably theory of mind ability (which would include behavioral rules) involves something like option 1, but could it also involve option 3? Minimal theory of mind cognition allows subjects to track others' intentions, perceptions, and beliefs *but not as such* (p. 28). However, this is ambiguous. It leaves open options 1, 2, and 3. Does minimal theory of mind ever or always involve representing mental states as such? I am not sure how B&A would answer these questions.

One thing that B&A could do to distinguish minimal theory of mind from behavioral rules is to argue that minimal theory of mind involves representing goals *as such* (option 2) whereas the behavioral rules approach involves only representing goals (option 1). So long as we retain B&A's non-mentalistic, non-representational definition of goal (the outcome to which a behavior is directed), option 1 is consistent with the behavioral rules approach.

This is one way using B&A's own framework that we can draw a contrast between the commitments of minimal theory of mind and the behavioral rules approach. However, B&A still would need to clarify the relevant cognitive difference between representing a goal and representing a goal as such. Because so much rests on this distinction, it is important to have a clear account of what it is. And in its current state, the account does not shed much light on this distinction.

#### **4. A Minor Objection**

B&A treat adult, infant, and non-human animal minimal theory of mind as roughly the same. Sections 1-3 use examples from adult, infant, and non-human animals as motivation for positing a minimal theory of mind. Moreover, B&A say, "Minimal theory of mind may be what enables those with limited cognitive resources or little conceptual sophistication, such as infants, chimpanzees, scrub-jays and human adults under load, to track others' perceptions, knowledge states and beliefs" (p. 28).

Equating these different cases of minimal theory of mind requires more defense. On the face of it, the minimal theory of mind cognition of scrub-jays and human adults under cognitive load is quite different. It would be helpful to know how B&A conceptualize the differences between, say, scrub-jay's, a toddler's, and an adult's minimal theory of mind. The discussion of the variations and extensions of principle

4 (p. 17) may offer the resources for different characterizations of adult, infant, and non-human minimal theory of mind.

As a final note, I should emphasize again that B&A's account is a needed and welcome addition to the dialectic. The critical attention on their account will, I hope, help bring into focus the elements that could be improved and perhaps pave the way for further middle-ground accounts. B&A's contribution to the literature will stimulate interesting and important debates about the nature of minimal theory of mind.