

Response to Westra

Shannon Spaulding

To cite this article: Shannon Spaulding (2020) Response to Westra, *Philosophical Psychology*, 33:6, 883-887, DOI: [10.1080/09515089.2020.1765325](https://doi.org/10.1080/09515089.2020.1765325)

To link to this article: <https://doi.org/10.1080/09515089.2020.1765325>



Published online: 13 May 2020.



Submit your article to this journal [↗](#)



Article views: 23



View related articles [↗](#)



View Crossmark data [↗](#)



Response to Westra

Shannon Spaulding

Department of Philosophy, Oklahoma State University, Stillwater, OK, USA

I am grateful to Evan Westra for his probing commentary on my book (Spaulding, 2018). His comments have spurred me to think much more carefully about my cynical analysis of our mindreading abilities. In his review, Westra raises an interesting and difficult methodological question about whether we can know how frequently we make mindreading errors. As he notes, my book focuses on mindreading errors partly as a corrective to an overly simplistic and optimistic view of our mindreading practices. I argue in the book that the standard story about mindreading vastly underestimates the complexity, diversity, and messiness of mindreading.

In an effort to correct this simple view of mindreading, I lay out various mindreading processes and detail the accuracy conditions for each of them. For example, when we are motivated by accuracy (or perhaps, more appropriately, *precision*, as Westra correctly and helpfully points out), we use deliberative mindreading strategies. Deliberative mindreading is likely to be accurate when we are (1) well practiced in this kind of deliberative mindreading, (2) not under extra cognitive load, (3) not skewed by motivational biases, and (4) basing our deliberation on good data, that is, data that is not limited or distorted (p. 45). When any of these conditions is not met, we are more likely to make mindreading errors. When our motivation for efficiency is stronger than our motivation for accuracy or precision, we employ various heuristics in mindreading a target. Which heuristics we use depend on judgments of perceived similarity to the target. These heuristics are more likely to be accurate when we (1) correctly diagnose our similarity or dissimilarity to the target, (2) we correctly adjust for situational context, (3) we correctly introspect our own mental states (when we are egocentrically projecting), and (4) we employ an appropriate stereotype (when we are stereotyping) (pp. 46–48). Again, when these conditions are not met, we are more likely to make mindreading errors. Finally, we sometimes have a variety of self-interested motivations which may influence how we mindread. These forms of motivated

reasoning may serve important psychological and social functions – for example, maintaining a healthy self-image and solidifying social bonds with in-group members, which is especially important in contexts of threat or competition – but they make mindreading errors more likely (p. 51).

My cynical analysis of our mindreading abilities is not a global skepticism. In general, I aim to describe the conditions in which certain mindreading strategies are likely to succeed or fail. Based on these conditions, the pessimistic conclusion I end up with is that when we are outside our small, typically homogenous circle of family and friends, we are likely to make mindreading mistakes.

The conditions for accuracy I describe are grounded in robust empirical research. However, Westra argues that we lack good data on the *frequency* of mindreading errors. Perhaps these conditions are often satisfied, or perhaps accuracy is less likely when these conditions are not met but, overall, still quite likely. Just because there are many ways to err in mindreading does not imply that we often *do* err in mindreading. We simply cannot tell from data on accuracy conditions how often we make mindreading mistakes. Westra points out that many of the tasks employed in the mindreading literature simply are not constructed to test the accuracy of ordinary mindreading. Furthermore, Westra highlights, empathic accuracy paradigms that *are* supposed to test the accuracy of mindreading problematically presuppose the accuracy of certain forms of introspection (which may turn out to be a form of first-person mindreading).

This is an astute observation. We do not have good direct evidence on the accuracy rates of mindreading. Although we lack direct evidence, we may be able to triangulate various measures to get a better picture of the accuracy of mindreading in various conditions. For instance, we can look for interpersonal agreement in the mindreading of a target. That is, do subjects looking at the same naturalistic social interaction attribute the same personality traits and mental states to targets? Do they make the same behavioral predictions? Are their predictions correct? It is one thing, as a spectator, to predict how a target will behave. It is quite another thing to predict interactive behavior. Thus, we can ask whether subjects accurately predict how a target will interact with them.¹

These measures do not entail accurate mindreading, of course. People may make the same mistaken inference, especially if various individuals are employing the same flawed stereotypes or are subject to similar types of distorting self-motivated reasoning. In such cases, we would see intersubjective agreement without accuracy. Moreover, one may inaccurately mindread a target and still make accurate behavioral predictions. This can happen in cases where the situational context strongly determines what an individual will do, regardless of what the individual is thinking or feeling. In this

kind of case, accurate behavioral prediction does not entail accurate mindreading.

However, if we can find cases of interpersonal agreement amongst individuals from different demographic backgrounds (individuals who are subject to different stereotypes, who have different motivations, and who display different in-grouping or out-grouping patterns), this would ease the concern about interpersonal agreement dissociating from accuracy. In cases where we see interpersonal agreement amongst demographically different individuals as well as accurate (individual and dyadic) behavioral predictions, it is plausible that the agreement and the accurate predictions are based on accurate mindreading. Finally, in order to get around the introspection confound that Westra highlights, one could have subjects watch actors depicting a naturalistic social interaction and ask subjects to infer the actors' mental states. This avoids the problem of targets having to introspect their previous mental states.

In order to test the accuracy of mindreading, we need to employ these tests in conditions that evidence suggests are likely to lead to error. For example, we need to interpret simulated social interactions and investigate their interpersonal agreement and predictive accuracy when subjects are under extra cognitive load, are likely to experience motivated reasoning (e.g., when they have something to lose or gain), when they have limited exposure to a target, when there are superficial demographic similarities or dissimilarities between the subject and the target, and so on. Combining the knowledge of when subjects are more likely to make mindreading errors with these tests will give us a better picture of the overall accuracy of our mindreading abilities. If subjects perform well² on the mindreading tasks even despite being in conditions less favorable to accurate mindreading, then this would give us good reason to think that our mindreading abilities are actually pretty good.

Though Westra is right that we lack data which directly bear on the frequency of mindreading errors, there are experimental paradigms that could be adapted to test this. We will have to wait for those data. In the meantime, to be a bit more speculative, we should, in general, expect our mindreading abilities to be adaptive. It would be surprising if our mindreading abilities were so bad that they were maladaptive. Typically, I do not go in for evolutionary just-so stories; however, if the readers will allow me just one intuitively plausible speculation, I venture to guess that the adaptive function of mindreading has more to do with promoting (individual and group) social wellbeing than accuracy per se. If this is right, then mindreading can be both adaptive as well as not particularly accurate in some contexts. I shall elaborate.

Although there is no agreed-upon timeline for when the capacity for mindreading evolved in human history, it is uncontroversial to say that it evolved in a time when humans lived in relatively small groups of similar and often genetically related people. This implies that mindreading evolved to be

useful in the context of small groups of similar people. Indeed, the empirical evidence suggests that we are less likely to make mindreading errors when we mindread familiar and similar people. In an evolutionary context of small, tight-knit social groups, outsiders would almost always represent competition and threat. It would make sense to treat people in your small, homogenous social group very differently from people outside that group. Thus, it is easy to see how in-group and out-group dynamics I describe in the book would emerge, for example, stereotypes, implicit and explicit biases, and dehumanization. The tendency to treat out-group members very differently has deleterious effects on the accuracy of mindreading, but it solidifies the social bonds of the in-group, especially in contexts of threat or competition. If this is right, then mindreading should be fairly accurate within small, homogenous groups but less accurate in large, diverse groups while still serving the adaptive function of promoting social wellbeing.

We no longer live in very small, homogenous groups of people, of course. Pretty recently in evolutionary history, we shifted from very small, homogeneous groups to large, geographically broad, and demographically diverse social circles. We now regularly encounter people of different races, nationalities, religions, cultural practices, expressive behaviors, and social norms. If my speculative story is right, then the mindreading skills that were adaptive in small, homogenous groups may be less accurate in contemporary social environments.

Thanks again to Evan Westra for his clear, insightful, and important commentary. His work is, as always, a pleasure to engage with. My hope is that this dialogue spurs more experimental and theoretical work on the accuracy of mindreading.

Notes

1. There are existing paradigms in the social psychological literature that ask these questions about traits, but we could adapt them to study mental-state inferences in general. See Zaki and Ochsner (2011) for an overview of these paradigms.
2. It is an open question how to define good performance. If we measure accuracy against the baseline of chance, as opposed to perfectly accurate performance, we will get a much more optimistic assessment of our mindreading abilities. This standard may seem too weak to some, however.

Disclosure statement

No potential conflict of interest was reported by the author.

References

- Spaulding, S. (2018). *How we understand others: Philosophy and social cognition*. Routledge.
- Zaki, J., & Ochsner, K. (2011). Reintegrating the study of accuracy into social cognition research. *Psychological Inquiry*, 22(3), 159–182. <https://doi.org/10.1080/1047840X.2011.551743>