

Reply to Schulkin

Jay Schulkin¹ has several reservations about simulation theory (ST) as an account of mind-reading; we have reservations about his reservations. Schulkin classifies ST as a modern variant of the argument from analogy (for other minds), but thinks this argument has long been refuted. Several comments are in order. First, not all simulation theorists construe the theory as a form of the argument from analogy². However, in our view, the resemblance is close enough to merit discussion³. We do not agree that traditional criticisms of the analogical argument render it hopeless for cognitive science. The chief philosophical criticism of it, in fact, concerns an epistemological question: could anyone be justified in believing in the mental states of others solely on the basis of their own mental states? This would be an induction from a single case, a very problematic induction. Whatever one's conclusion about this epistemological issue, it does not concern the cognitive science of mind-reading. Cognitive science is interested in the processes by which children and adults actually arrive at third-person mental-state attributions, whether or not these processes yield 'justified' beliefs. The evidential warrant for interpreters' attributions is beside the point, because there is no guarantee that the basic cognitive mechanisms nature has put in place use evidentially impeccable inferences.

Schulkin criticizes the analogical approach on the grounds that 'we know a lot about the world and other people's experiences even when there is no corresponding experience in us'. In offering this as a criticism, he seems to assume that ST is committed to the notion that each and every instance of mind-reading involves a simulation heuristic. That is not the version of the approach that we take (see Ref. 3). People often engage in mind-reading based on rules of thumb that may be derived from earlier uses of simulation, but do not require current simulation. This is compatible with the idea that we use our own mentality as the 'home base' for interpreting others. Our folk psychological knowledge of the mental states of others ultimately rests on our own experience even if it doesn't mean that every instance of new mentalistic knowledge of others requires a corresponding experience in ourselves. It is noteworthy that even some psychologists who otherwise characterize themselves as theory-theorists (e.g. Andrew Meltzoff) explicitly appeal to analogical inferences: "When infants see others acting "like me", sharing their behavioral states, they project that others are having the same mental experience that is concomitant with those states.... [I]nfants infer intentionality in others based on an analogy to the self."⁴ So the analogical inference idea is not a dead

letter within cognitive science, even outside the simulation-theory camp.

Schulkin contends that simulation is a computationally more laborious process than theoretical inference, especially when the target's mental state is quite different from that of the attributor. This contention needs much more careful spelling out. The crux of ST is that an attributor attributes states to a target by pretending to be in his initial situation and then running a simulation heuristic to determine further states. Is pretending to be in a very different state from one's current state computationally more difficult or time-consuming than pretending to be in a similar state? That isn't obvious. Nor is it obvious that dealing with different states is *not* in fact more difficult and time-consuming for attributors. That is an empirical question. So even if ST has the implication that Schulkin claims, that is not necessarily a flaw in ST.

On the topic of computational complexity, theory-theory (TT) seems to have a greater burden than ST. TT posits a body of generalizations known by the attributor. In an attributional task, the attributor must posit the target's initial states, access the appropriate laws, and then apply those laws to the initial states to infer further states. ST avoids the need for knowing and accessing generalizations because it posits the use of one's own cognitive operations. These operations can just be *used* without the attributor necessarily knowing and accessing their lawful properties. Finally, because of the holism involved in standard forms of TT, the meanings of one's mental-state concepts will be constantly changing as one acquires new laws (in terms of which the concepts are defined). The constant updating of the meanings of mental-state concepts adds a further computational burden.

Schulkin likes the TT story because he maintains that the contents of the mind are as theoretical as the contents of the kidney. Does he mean that both are unobservable? The claim that mental states are unobservable is one that many theory theorists make. But is it correct? Certainly it is true that someone else's mental states are as unobservable to an attributor as someone else's kidneys. But mental states in general may not be unobservable if it is possible to use inner 'observation' of one's own mental states and learn properties of these states with the help of such 'observations'. The idea of internal observation, or self-monitoring, continues to be a widely accepted approach, if not a consensus approach, among both philosophers of mind and cognitive scientists⁵⁻⁸. This approach should not be tarred with the brush of Cartesianism. Internal monitoring is not committed to Cartesian infallibility or indubitability. Note, moreover, that ST does not necessarily require conscious processing or

the use of introspection. Mirror neuron activity, for example, is presumed to be automatic and unconscious. Our paper⁹ did not explicitly invoke introspection at any point. However, we do think that mature, folk-psychological simulation is sometimes conscious and sometimes employs introspection.

If the mental states posited by folk psychology were just as unobservable as the states of the kidney, they should never be parts of our conscious experience. Folk psychology ought to posit states that are wholly unobserved in the sense that their instances are never observed by external perception or internal monitoring. At least some of the states of folk psychology should be like this. The fact of the matter is quite different however. The stock inventory of folk psychological states – seeing, believing, wanting, intending, fearing, and so forth – are all states *some* of whose tokens are conscious, and hence introspectively accessible. This is not what one would expect if the basic cognitive operation of mentalistic concept formation paralleled that of theoretical science.

Schulkin spends a lot of time discussing the term 'detached' – this is unfortunate. The word occurred once in our article, and as readers familiar with the debate will know, it does not play a central role in this debate. Perhaps a better articulation of the ST/TT contrast (a contrast more fully illustrated by our original Fig. 3; see Ref. 9) is to say that ST postulates 'shared' states in mind-reading, whereas TT postulates no such sharing. According to ST, when I 'mind-read' your intention state I form a pretend intention with the very same content. According to (pure forms of) TT, however, all mind-reading involves purely credal or belief-like states, for example, beliefs *about* intentions but not themselves intentions. This distinction implies no concrete/abstract distinction, and no up-close/distant distinction, matters discussed at length by Schulkin but which have little bearing on the real issues.

The idea of 'shared' states is the focal point of our interest in mirror neurons (MNs). In observationally stimulated MN activity, the observer shares a certain state with that of the target actor. If this sort of mechanism is the basis for, or a precursor of, interpersonal mind-reading, then this suggests a process of mind-reading that is more like simulation than like theorizing. It is noteworthy that mirroring activity does not take place only in the premotor cortex. A similar phenomenon is apparently found in pain-related neurons. Hutchison *et al.* have studied pain-related neurons in the human cingulate cortex¹⁰. Cingulotomy procedures for the treatment of psychiatric disease provided an opportunity to examine whether neurons in the anterior cingulate cortex of awake humans

respond to painful stimuli. Serendipitously, it was noticed that a cell that responded to noxious mechanical stimulation, but not to a noxious heat stimulus, also responded when the patient merely watched pinpricks being applied to the examiner's fingers.

In fact, we posit that MNs could be just one instance of a much more general matching mechanism that uses internal representations of goals, emotions, body states and the like to map the same states in other individuals. Following the suggestion of Adolphs that the somatosensory-related cortices of the right brain might be relevant for social cognition¹¹, one could speculate that this is so because of the presence of 'somatosensory MNs' that allow the observer to map other individuals' body parts on his/her own body parts. A possible suggestion, therefore, is that by means of such diversified matching systems, the observer is able to 'recognize' other individuals as his/her social partners. Future experiments will have to be developed to test this hypothesis.

We are completely puzzled by Schulkin's discussion of the developmental literature on desires and beliefs. We do share his doubts about certain conclusions that have been

drawn in this literature, especially about non-representational intentionality¹². But we do not understand at all how this part of his discussion is supposed to constitute an argument or consideration in favor of TT over ST. In general, his arguments leave us quite unmoved from our previous conclusions.

Alvin Goldman

Department of Philosophy, University of Arizona, Tucson, AZ 85721-0027, USA.
tel: +1 520 621 3120
fax: +1 520 621 9559
e-mail: goldman@u.arizona.edu

Vittorio Gallese

Instituto di Fisiologia Umana, Università di Parma, Via Volturno 39, I-43100 Parma, Italy.

References

- 1 Schulkin, J. (2000) Theory of mind and mirroring neurons. *Trends Cognit. Sci.* 4, 252–254
- 2 Gordon, R. (1995) Simulation without introspection or inference from me to you. In *Mental Simulation* (Davies, M. and Stone, T., eds), pp. 53–67, Blackwell
- 3 Goldman, A. (1989) Interpretation psychologized. *Mind Lang.* 4, 161–185

- 4 Meltzoff, A. and Brooks, R. 'Like me' as a building block for understanding other minds: bodily acts, attention, and intention. In *Understanding Intentions and Intentionality: Foundations of Social Cognition* (Malle, B. et al., eds), MIT Press (in press)
- 5 Armstrong, D.M. (1980) What is consciousness? In *The Nature of Mind and Other Essays*, Cornell University Press
- 6 Lycan, W.G. (1996) *Consciousness and Experience*, MIT Press
- 7 Baars, B.J. (1988) *A Cognitive Theory of Consciousness*, Cambridge University Press
- 8 Goldman, A. (2000) Can science know when you're conscious? Epistemological foundations of consciousness research. *J. Conscious. Stud.* 7 (5), 3–22
- 9 Gallese, V. and Goldman, A. (1998) Mirror neurons and the simulation theory of mind-reading. *Trends Cognit. Sci.* 2, 493–501
- 10 Hutchison, W.D. et al. (1999) Pain-related neurons in the human cingulate cortex. *Nat. Neurosci.* 2, 403–405
- 11 Adolphs, R. (1999) Social cognition and the human brain. *Trends Cognit. Sci.* 3, 469–479
- 12 Goldman, A. Desire, intention, and the simulation theory. In *Understanding Intentions and Intentionality: Foundations of Social Cognition* (Malle, B. et al., eds), MIT Press (in press)

Monitor

Summaries of recently published papers of interest to cognitive scientists. Readers who would like to contribute to this section, by identifying appropriate papers and writing short summaries, should contact the Editor (tics@current-trends.com).

Irregular transcripts?

One of the most famous phenomena in language acquisition is the over-regularization of the past tense of English. According to conventional wisdom, children's use of the English past tense through development follows a U-shaped curve: initially, children use all past forms correctly; then, correct performance dips as children apply the regular past ending ('-ed') to irregular stems, creating over-regularized forms such as 'goed' and 'singed'; finally, children recover from this error and produce all forms correctly. The meaning of this pattern was hotly debated – does the dip reflect the over-application of a newly learned rule, or is it a natural by-product of a connectionist association network? – until Marcus *et al.* pulled the rug out from under the controversy¹. After a close examination of transcripts of children's speech, they found that actual rates of over-regularization were very low (around 5%) and hardly worthy of special analysis at all. Marcus *et al.* argued that these small error rates are simply the result of occasional memory retrieval difficulties with the irregular past forms,

which normally serve to block the regular forms.

In a recent paper, Maratsos returned to the data in an effort to reinstate the conventional wisdom that there is a real stage in development in which both the correct irregular forms and the over-regularized past forms are in competition in the child's grammar². Maratsos argues that the transcripts, which represent at best only a few hours of speech per child per week, cannot be used directly as a means to measure the time course of acquisition. According to Maratsos, the acquisition of memorized forms (such as irregular past forms) is most probably token-dependent, and the vast majority of the tokens of each verb type a child hears (and produces) will occur outside of the available transcripts. Given some plausible estimates of how many tokens of an irregular past verb a child needs to hear in order to move beyond over-regularization, Maratsos shows that for high-frequency verbs, the over-regularization stage would last only a matter of weeks and therefore be largely invisible on the transcripts. Only for relatively infrequent verbs will the over-regularization period

last long enough to be detectable in the transcripts. Maratsos undertook a verb-by-verb analysis of past tense forms and showed that indeed, for relatively infrequent verbs, over-regularization rates were quite high (on the order of 30 to 50%). This paper goes a long way towards re-establishing the legitimacy of the past tense over-regularization phenomenon, but more than that, it raises important issues of how best to use the extensive transcripts of child speech now widely available. These transcripts have had a profound influence on how language acquisition research is conducted; Maratsos warns us that this rich data source must be accompanied by equally rich theories of how to interpret it.

References

- 1 Marcus, G. et al. (1992) Over-regularisations in language acquisition. *Monogr. Soc. Res. Child Dev.* No. 228
- 2 Maratsos, M. (2000) More over-regularizations after all: new data and discussion on Marcus, Pinker, Ullman, Hollander, Rosen and Xu. *J. Child Lang.* 27, 183–212

Students

Did you know you are entitled to a 50% discount on a personal subscription to *Trends in Cognitive Sciences*? Send off the bound-in subscription card now!