

(an IR, in Goldman's terms), one must also consider the relation between this IR and some other entity. Contrary to Gopnik's claim, it is the relation of one's psychological state to the content of that state, and the relation of the content to the world, and our idea of the world (considered as a thing in itself – e.g., as the thing that gives rise to our perceptions), that are hypothetical entities, not one's own direct awareness of the contents of one's mental states. The latter awareness is directly given – in fact, it is the only thing of which one can ever be directly aware: The rest is *noumenal*.

Gopnik's instantiation of the psychologist's fallacy most likely contributes to her misreading of Searle (1990). Searle claims that states that are not at least in principle accessible to consciousness lack intrinsic intentionality, and Gopnik takes this to mean that "the intentionality of psychological states is known directly in the first-person case" (sect. 2, para. 5). Conscious brain states surely *are* intentional (Zelazo & Reznick 1990); it is a different thing altogether to claim that their intentionality is *known*.

So, 3-year-olds may be directly aware of the content of their mental states. And these mental states are necessarily *about* something, and so are intentional (Brentano 1973). When 3-year-olds report that they think there are pencils in the box, they directly access and report the contents of a mental state. What they do not understand is that their mental states have content, and that there is some *other* reality (defined in terms of the noumenal world or in terms of other people's mental states), that the content of their mental state is a representation of. This is an understanding of a higher-order relation among terms (some of which are hypothetical) that is hypothetical, and it is a late development. The logical structure of tasks that assess this understanding is the same for tasks that require reasoning about one's own mental states and about other people's mental states. As such, it is not surprising that performance on these tasks is correlated.

There are differences, however, between reporting on one's mental states and reporting on other people's mental states. When one reports on another's mental states, one clearly does not have direct access to those states. Contra Goldman, one uses knowledge of functional relations to infer those states. Goldman is right that one may rely on simulation and on analogy to one's own IRs to understand another's mental states. However, simulation and analogy plainly rest on functional information. How does one know which situation to put oneself in to simulate another's mental state? One must identify the relevant input-internal state-output relations and imagine oneself as participating in those relations. The need for functional information becomes clearer when one cannot duplicate another's situation exactly, as when one says, "I have never experienced being discriminated against on the basis of race, but it may be something like . . .," drawing an analogy to a similar (*in the relevant functional respects*) situation. Direct access to one's own past mental states is necessary here but so is functional information and a certain level of logical sophistication.

A complementary case may be made for self-attribution. (Note that Goldman's account must apply to the self-attribution of past mental states: Introspection is really retrospection, because the moment of which we speak is already past.) Although we have direct access to some of our past mental states, access to other of our mental states would seem to rely on the use of functional information to set up the conditions for recall. This is analogous to setting up the proper simulation of another's mental states. Accessing the appropriate memory in representational change tasks requires (1) being able to handle embedded-rule reasoning, (2) using the proper functional information to access just the relevant memories, and (3) being able to exhibit a certain amount of representational flexibility (if the content of the present mental state is more salient, children may focus on this). In some cases, functional information overrides phenomenological information so that one may conclude that one must

have thought one thing, even though one remembers thinking another.

EDITORIAL COMMENTARY

One can argue about whether or not I have a "what it's like to believe that the cat is on the mat" quale and whether or not having that belief is uniquely accompanied and identified by having that quale (I happen to think it is). But it seems much less disputable that that belief must at least be *grounded* in having qualia for "cat" and "mat," for otherwise it would be hanging from a contentless formal skyhook. Now it might be that as a matter of fact its actual and potential causal interactions with the world ground that formal skyhook functionally (I happen to think that's true too), but unless that functional grounding also gives it qualitative content, it is not clear why the belief should be regarded as *mental* at all. Purely functional "beliefs" can be had by coat hangers, too. Or, to put it another way, surely it cannot be a mere coincidence that the only creatures that have beliefs are also the only creatures that have qualia, and that qualia seem to be the only thing that puts mental flesh on the formal/functional bones of what those beliefs are about.

Author's Response

Theories and illusions

Alison Gopnik

Department of Psychology, University of California, Berkeley, CA 94720
Electronic mail: gopnik@cmsa.berkeley.edu

R1. Theories and illusions. My commentators raise a wide variety of issues ranging from highly empirical questions about the details of various experiments to highly abstract philosophical concerns. I have organized these into several general topics, hoping not to minimize the subtlety of particular arguments. I have tried to deal first with a number of technical questions about the experiments, many of them from other developmental psychologists. Then I have moved on to some more abstract conceptual and definitional questions and philosophical arguments. Finally, in the last sections I have tackled what seem to me the most interesting questions, substantive theoretical ones about the nature and origins of our understanding of the mind.

R2. Information-processing alternatives redux. At least one prediction of the target article was completely confirmed: Many commentators have alternative information-processing-like explanations for the children's difficulty with the false-belief task. These include Butterworth, Campbell & Bickhard, Dittrich & Lea, Goldman, Gurd & Marshall, Josephson, Leslie et al., Plaut & Karmiloff-Smith, Russell, Siegal, Zaitchik & Samet, and Zelazo & Frye. I am not too perturbed by this, however, for two reasons. First, all these commentators have *different* alternative explanations for the false-belief task. If there is an information-processing alternative, it certainly is not obvious. In contrast, the "theory-theory" accounts advanced by myself and others (such as Flavell [1988],