Knowing levels and the child's understanding of mind

Robert L. Campbell[a] and Mark H. Bickhard[b]

[a]*Department of Psychology, Clemson University, Clemson, SC 29634-1511 and* [b]*Departments of Philosophy and Psychology, Lehigh University, Bethlehem, PA 18015*

Electronic mail: [a]*campber@clemson.bitnet and* [b]*mhb0@lehigh.edu*

[Gol, Gop] We are pleased to see that the developmental study of children's conceptions of mind is beginning to have an impact on philosophy of mind. After all, if we take the program of genetic epistemology seriously, that is how things are supposed to work. Gopnik presents a solid summary of her own research, and studies by Flavell, Wimmer, Perner, Wellman, and others, showing that there is a major transition in children's understanding of their own and other people's minds around age 4. Before this transition, children have trouble understanding their own and other people's false beliefs, deceiving other people with the express goal of getting them to accept false beliefs, and differentiating the way an object looks from the way it actually is. Gopnik, like a number of other investigators, proposes that children are developing a "metarepresentational" capacity to think about their own and other people's mental representations. Like most of these other investigators, she interprets this as a representational *theory* of mind.

From an interactivist standpoint, we would agree that a metarepresentational capacity develops around age 4, but we do not cast that capacity as a theory. Instead, we draw on a conception of levels of knowing (Bickhard 1973; 1978; 1980; Campbell & Bickhard 1986) whose basic intuition is as follows: Knowing is accomplished by a goal-directed system interacting with an environment. Knowing is irreflexive; the system can know properties of the environment by interacting with it, but it cannot know anything about itself, even though some of its own properties might be useful to know. A subsystem that interacts with the knowing system, much as the knowing system interacts with the environment, however, could know these properties. Specifically, a second-level system could know and learn about the first-level system by interacting with it. Once the hierarchy gets started, a third-level system could know things about the second-level system; a fourth-level system could know about the third, and so on, unboundedly.

According to interactivism, reflexive consciousness requires a second knowing level to interact with the first. In that manner, it is possible to acquire knowledge about knowledge and belief. The hierarchy of knowing levels is also implicated in development through stages; instead of being defined in terms of characteristic mathematical structures like the Piagetian stages, knowing-level stages are defined in terms of the level of knowing at which knowledge is being constructed. The process that is responsible for ascension to the next knowing level we call reflective abstraction, to borrow a term from Piaget (1977; 1986).

Understanding false beliefs, engaging in deliberate deception, sorting out appearance from reality – and very likely other developments not mentioned by Gopnik, like the emergence of autobiographical memory (Nelson 1992) – are all instances of the onset of level 2. So, we would claim, are changes in causal and classification reasoning that require reflective abstraction on prior understandings, but not thought about belief as such (Campbell 1992; Campbell & Bickhard 1986). All these changes begin around age 4.

The knowing-levels conception also has an important bearing on the role of maturation in the changes that happen around age 4. Gopnik states

> One possibility might be that the 3- to 4-year-old shift is the result of the maturation of an innately determined capacity . . . . So far as I know no one actively working in the field, not even Leslie, has suggested that the 3- to 4-year-old shift is the result of the maturation of such a module. (note 8)

From an interactivist standpoint, it is only possible to ascend to knowing level 2 if a physically differentiated subsystem is present that can interact with the level 1 knowing system. (Once level 2 knowing is possible, higher levels can be reached on a purely functional basis, without extra hardware.) An empirical consideration of the earliest possible signs of level 2 knowing in human development led Bickhard (1973; 1978) to place the earliest transition at age 4; consequently, interactivism predicts the maturation of the physical second-level knowing system (not a Fodorian module!) around this age. Obviously, the maturational hypothesis needs to be investigated at a neural level; we have been proposing and elaborating this idea for nearly twenty years (Bickhard 1973; 1978; 1980; 1992; Campbell 1992; Campbell & Bickhard 1986).

Now on to some secondary themes. Gopnik, like Perner (1991b), Wellman (1990), and apparently most others in this line of work, attributes to young children a "theory" of mind (sect. 6). The reasons for doing so have never appealed to us. As Goldman rightly points out (in his sect. 10), such talk of theories is extremely loose, ignoring, for example, any reference to nomological or lawful generalizations. Indeed, all that Carey (1985), Murphy and Medin (1985), Keil (1989), Perner (1991b), Wellman (1990), and others in this camp seem to require of theories is that they be coherent networks of concepts that can be used for predicting and explaining. This formulation borders on the vacuous; given what we know about human knowledge, even the limited sorts possessed by newborn babies, is there any of it that *wouldn't* qualify as a "theory"? Would a Piagetian scheme not qualify? A set of production rules? (In treating "information-processing (IP) alternatives" in sect. 5 as mere "performance" considerations, Gopnik, perhaps unwittingly, trivializes IP theory.) The only thing that would not qualify as a theory is a collection of atomistic, self-encapsulated, encoded concepts. From the interactivist standpoint, knowledge simply cannot take that form (Campbell 1992). For Goldman, however, it can, and he unwittingly lends credence to the "theory-theory" by presenting as his alternative an unorganized aggregate of mental state concepts. Our own preference is to follow those who study the use, modification, and testing of explicit hypotheses and theories later in development (notably Kuhn et al. 1988) and restrict talk of theories to that arena.

We must also comment on Gopnik's peculiar and apparently distorted invocations of Gibsonian theory (sect. 3.2 and note 5), including her claim that 3-year-olds, who cannot yet think about representation, must therefore have a "Gibsonian" conception of the relation of mind to world. In her assertion that "the relation between real things in the world and our perception of them is a direct causal link, almost a transference," she is conflating Gibson (1966; 1977; 1979) with Dretske (1981). There is a similar conflation with Dretske in her assertion that a Gibsonian approach cannot deal with error in the form of false beliefs or misleading appearances; were this the case, why, for instance, did Gibson try to explain visual illusions?

More misunderstandings of Gibson and of perception show up in Gopnik's claim that our belief, as adults, in the special and privileged nature of our first-person knowledge of ourselves, is an "illusion of expertise" (sect. 7). Gopnik asserts that when a chess master claims to see the forces threatening a king, this cannot be real perception, because that would have to involve an

> experience [that is] reliably, and reasonably directly, caused by the object. . . . In developing forms of expertise, we construct an implicit theory of the realm in which we are expert. Various kinds of genuine perception act as important evidence for that theory. . . . Given this [genuinely perceptual] evidence, or even a single piece of it, the diagnostician draws on vast, nonperceptual, theoretical knowledge to make implicit inferences about the patient. He quite appropriately applies the theory, "the patient has cancer" . . . . [But] from his first-person view, the cancer may simply be perceived.

Note again the emptiness of talk about implicit theories – what kinds of knowledge would Gopnik exclude from counting as theory? Moreover, this argument ignores the discussion of genuine, extended kinds of perception in Gibson. Indeed, Gopnik's strictures seem to rule out any and all perceptual learning! When chicken-sexers think they can see whether a newly hatched chick is male or female, is this just another "illusion of expertise"? At a deeper level lie questions like: What does it mean to say that perception is direct? Could perception be mediated by encodings? In what sense does perception involve inference? Gopnik just assumes that theory involves mediating encodings and perception does not (for more about these issues, see Bickhard & Richie 1983).

Gopnik has the advantage that her ideas are grounded in a solid, fairly well elaborated research program in developmental psychology. By contrast, we find Goldman's own research program to be considerably less promising. His convoluted investigations of functionalism and folk psychology strike us as largely beside the point, because they all rest on an inadequate model of categorization. In his section 2, he proposes that people ascribe mental states to themselves and to others by matching category representations (CRs) for mental state words to instance representations (IRs) of mental states (or, in his preferred alternative, to mental state instances as such). "The content of such an IR will be something like 'A current state (of mine) has features $\phi_1, \ldots, \phi_n$.' Such an IR will match a CR having the content: $\phi_1, \ldots, \phi_n$. Our aim is to discover, for each mental word *M*, its associated CR. . . ." Now the recourse to matching is widespread in cognitive science; conventional information-processing models like those of Anderson (1983) and Newell (1990) depend on it, because production rules fire only when a match is detected between their symbolic initial conditions and symbols present in working memory.

But matching is completely useless for explaining categorization or pattern recognition. Matching models raise the same epistemological questions all over again in microcosm. If pattern recognition has to be explained by the matching of component features in encoded representations, is not the matching of the features itself just as badly in need of explanation? How does ϕ_1 in a CR come to represent "the same" ϕ_1 in an IR? By parity of argument, this would require us to match up subfeatures ϕ_{11} through ϕ_{1k} between ϕ_1 in the CR and ϕ_1 in the IR; those, in turn, would embroil us in the matching of subsubfeatures; and so on ad infinitum. This is, in any case, a style of argument that should be familiar to contemporary philosophers of mind, most notably via Kripke's (1982) interpretation of the later Wittgenstein.

The seduction here is that, since ϕ_{1cr} is in the same notation as ϕ_{1ir}, it would seem that this match could be performed purely functionally – by some sort of direct comparison, perhaps. But nothing in a CR is *identical* to things in an IR, so this match must be a judgment of similarity or identity in some respects, and it is this sort of judgment that was to be explained in the first place. Note that if judgments of match were naturalistically cashable this easily then the general problem of intentionality would be enormously simpler, if not already solved.

Besides, if connectionists have accomplished anything at all in recent years, they have shown how patterns can be differentiated without reliance on symbols or matching. Yet Goldman pays no heed to connectionism whatsoever. Interactivism (Bickhard 1973; 1980) offers an alternative as well, though we lack the space to expound it here.

In sum, ongoing research about children's understanding of mind has proven instructive and challenging for psychology and philosophy alike, and we hope it will prosper. At the same time, we hope that more attention will be devoted to thinking about the deep and sometimes slippery epistemological issues that are bound up with this kind of research. With the exception of Perner (1991), workers in this field have displayed more enthusiasm in designing and running empirical studies than care in thinking about their implications.