REPLY

GROWING UP APE

Daniel J. Povinelli

Over 2 years have now elapsed since Timothy Eddy and I completed the series of experiments reported in this *Monograph*. Meanwhile, our seven apes have grown 2 years older. This passage of time has allowed us to conduct some additional research on these same animals that has direct implications for the important theoretical points raised by both Hobson and Tomasello as well as for some of the hypotheses we entertained at the end of Chapter VI. In particular, longitudinal follow-ups using some of the same tasks have provided us with some intriguing new information that further constrains the interpretations of our data set. In this Reply, I draw on this new work to address two recurring and interrelated themes in both our *Monograph* and the Commentaries: (1) the relation between experience and conceptual understanding in theory of mind development and (2) the relation between performances based on behavioral/procedural rules and those that are augmented by an additional, conceptual understanding of mental states. Finally, I use this discussion to sketch an alternative view that could account for the confusing mosaic of similarities and differences we see when comparing the behavioral and cognitive systems of humans and apes.

Is Experience the Master?

Amount of Experience

Anyone reading this *Monograph* will note our hesitation in automatically generalizing our findings to adult chimpanzees—a concern that Tomasello underscores in his Commentary. Indeed, one of the alternative hypotheses that we outlined in Chapter VI was that although 5–6-year-old chimpanzees

do not seem to understand seeing-as-attention, older chimpanzees might. To put it crudely, apes grow up, too. Readers will also note our insistence that although the pattern of our subjects' performance was best predicted by a learning framework, the very fact that in the end they were quite adept at choosing the experimenter who could see them—no matter how we posed the question—left us in a difficult quandary. On the one hand, it was possible (indeed likely) that with experience our apes had learned a set of procedural rules that were sufficient to solve our tasks, without any concomitant understanding of attention or "seeing" per se. On the other hand, even though the emergence of these rules can be easily explained in terms of general learning theory, it is also possible that this experience simultaneously endowed our apes with a genuine understanding of seeing-as-attention. As we noted, choosing between these two alternatives is not easy. More important, either choice has implications for our interpretation of similar research with young children. For example, if our apes came to provide the appearance (but not the reality) of a mentalistic understanding of seeing, who is to say that the children we tested were not in the same position? To be sure, the degree to which apes or children are able to generalize to new situations may give us some indication of what they have really learned, but a moment's reflection will reveal that there are ways in which such generalization could occur without an attendant conceptual theory of seeing.

Tomasello misinterprets our raising this warning flag as meaning that we believe that very young children may not really possess theory of mind skills. To the contrary, I have been convinced for quite some time by the same body of evidence to which Tomasello refers (some of which he has been instrumental in collecting) that by 18–24 months human toddlers have taken their first firm steps into a psychological arena in which the social milieu is understood in terms of private mental states (for an early statement of this position, see Gallup & Suarez, 1986). But understanding the conglomerate of changes that occur at this general developmental period (18–24 months) can tell us only so much about the underlying cognitive processes supporting any given behavior that emerges during (or after) this period. As a case in point, many commentators (including Tomasello) have expressed doubt about the cognitive significance of the ability of many chimpanzees to recognize themselves in mirrors. These criticisms are worth consideration, but the ability to recognize oneself in mirrors typically emerges in human infants at about 18–24 months (Amsterdam, 1972; Lewis & Brooks-Gunn, 1979). So in this case the fact that a skill emerges alongside a cluster of others is not a necessary indication of its underlying psychological cause (for data concerning the correlation between the emergence of self-recognition in mirrors and other behaviors, see Asendorpf & Baudonniere, 1993; Bischof-Köhler, 1988, 1994; Johnson, 1982; Lewis, Sullivan,

Stanger, & Weiss, 1989). Likewise, at an even more detailed level of analysis, Gopnik and Meltzoff (1986) have provided theoretical and empirical grounds for suspecting that many of the cognitive achievements during this period may be occurring in relatively separate domains.

Thus, the fact that young 2-year-olds have some understanding of mental states such as desire and attention, but that 2½-year-olds did poorly on some versions of our test of seeing-as-attention, does not uniquely call the validity of our task into question. Furthermore, it says little one way or the other about how competence on such tasks is constructed in the first place. Clearly, our question remains as salient now as it was 2 years ago: How do we distinguish between competence on theory of mind tasks that is the result of procedural rules derived from sheer experience versus competence that is both derived from experience and accompanied by conceptual understanding? Transfer tests of the type that we have employed here and elsewhere can help, but the manner in which we have constructed the problem may leave one wondering if we are left forever chasing a moving target.

Fortunately, one way of distinguishing between these alternatives serendipitously presented itself to us 13 months after the completion of the final study with our apes that we reported in this *Monograph*. In the context of a very different set of studies, we once again administered several of the seeing-versus-not-seeing treatments to our subjects. Initially, our protocol merely called for administering the eyes-open-versus-closed (C′′′′) treatment to show that our apes were still sensitive to this distinction and would respond in the same fashion as they had in the studies reported in this *Monograph*. Much to our surprise, however, the animals no longer seemed to have a preference for the experimenter who had his or her eyes open (Fig. R1). Even after 48 blocked trials of this treatment using the same stimulus configurations outlined in Experiment 13 (Chap. IV), our seven apes were responding at chance levels. Puzzled by this finding, we turned to the screens-versus-no-screens (B′′) treatment because we believed that it was more visually salient and because our apes had had considerably greater experience with it than with the eyes-open-versus-closed treatment. Again, however, the animals performed randomly—only gradually showing evidence of learning across the 12 probe trials they received (Fig. R1). Stepping back, we decided to administer four probe trials of the back-versus-front (C) treatment. After all, this was a treatment that the subjects had "understood" immediately nearly 2 years earlier. And, just as we expected, the animals' performance shot up to near perfect levels (Fig. R1).

The results of this longitudinal follow-up can be interpreted in a number of ways, but in the interest of space let me focus on just two. First, a strict learning theorist has a ready-made explanation of these findings: in the course of participating in a large number of additional studies in the intervening 13 months, our subjects had long since abandoned or forgotten

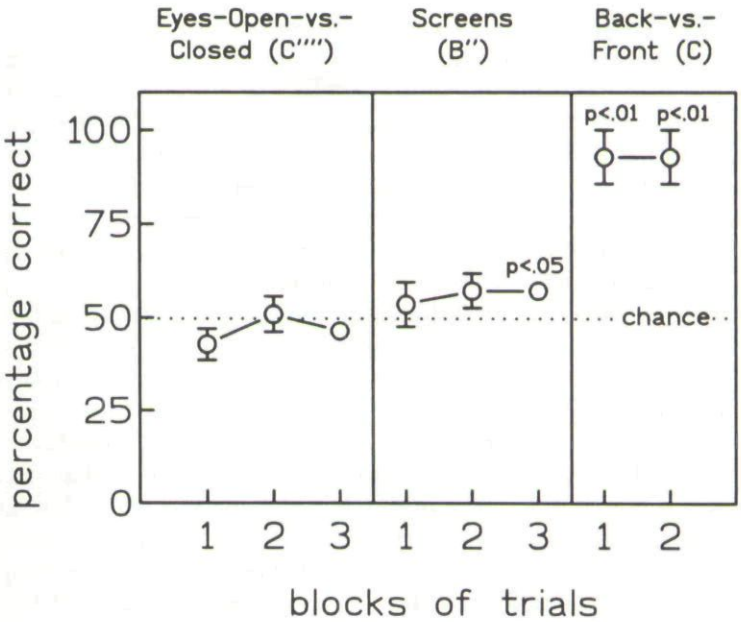

FIG. R1.—The results of a 13-month longitudinal follow-up on seven chimpanzees' performance on three of the experimental treatments used in Experiments 1–14. The subjects' mean percentage correct (± SEM) is shown in blocks of trials (C'''' in blocks of 16 trials, B'' in blocks of four trials, C in blocks of two trials). Trials were administered using the same general methods as described in Chaps. III and IV.

their old rules, having updated, revised, and/or replaced them numerous times. In contrast, their immediate and excellent performance on the back-versus-front (C) trials indicates their retention of a very strong rule or social disposition—a rule or disposition that continued to be useful in many of the additional studies in which they had participated. (For example, although only one experimenter had been present in most of these intervening studies, the apes were nonetheless required to orient to this person. Therefore, being attracted in general to the frontal stimulus of a human was still a very useful behavioral rule.)

I am uneasy about arguing too strongly in favor of this narrowly focused learning interpretation because its implications are so far-reaching. For example, if this explanation is correct, it means that our apes' generally successful performance at the end of the studies reported in Chapters III and IV can be clearly and conceptually distinguished from the successful performance of the young children we tested. Even if our young children had received numerous semistructured experiences with instances of "seeing" and "not seeing" prior to participating in our tests (so many in fact that their first trials on our task might be more equivalent to the apes' final

trials), they may have nonetheless parted conceptual company with the apes early on. For example, imagine these same children returning to our laboratory a year after participating in our tests and being retested on some new and arbitrary test of seeing-as-attention. The outcome cannot be in doubt. In the course of their development, through some combination of experience and hardwiring, they have constructed an enduring, fundamental understanding of the social universe—a system of understanding other minds (and maybe even the world as a whole) that transcends specific rules needed to navigate fairly narrow regions of ecological space, a system of understanding that we might truly wish to adorn with the label *theory* (see Gopnik & Meltzhoff, in press). But not, apparently, our apes. Although they are quite clever, our chimpanzees' striking difficulty on their longitudinal follow-ups provides learning theorists with additional support for their interpretations. The apes appear tied to social dispositions and procedural rules that are forged by epigenetic interactions, executed with precision, generalized to similar situations, but never integrated as part of a broad, interpretive apparatus with which to understand the minds of those around them.

Yet the learning theorist is not alone in vying for our conceptual allegiance; other interpretations of these longitudinal data are available. For example, we might imagine that when it comes to the development of theory of mind, experience is less the master than Alison Gopnik and others have speculated (e.g., Gopnik, 1993; Gopnik & Meltzoff, in press). Instead, the epigenetic pathways controlling the expression of theory of mind abilities in humans (and possibly in other species as well) may be tightly canalized. Across a fairly wide range of environments and experiences, the exact timing of the development of their cognitive capacities may be fairly uniform. The timing of the apparent conceptual transitions detected during infancy and young childhood may be real and may reflect maturational changes in neural systems subserving theory of mind skills.

However, a thorough knowledge of the rate and timing of human development in theory of mind will not necessarily guarantee an accurate understanding of the rate and timing of chimpanzee development. Indeed, there is some evidence that certain general aspects of cognitive development are temporally compressed in humans, relative to chimpanzees (see Povinelli, 1994, 1996). From this perspective, the longitudinal data reported in Figure R1 could simply mean that we should extend upward to about 7 years the lower age limit at which apes might develop a theory of seeing-as-attention. In our present state of ignorance, we should seriously entertain this possibility. But if we continue to obtain similar results as those reported here as these apes' adulthood approaches (and as we begin to work directly with adult apes), at some point we will be forced to abandon this developmental explanation of their differences from young children. Indeed, as most of our apes' eighth birthdays are approaching, they continue to re-

spond to even our simplest tests of an appreciation of attention in ways best predicted by a learning framework (see Povinelli, Bierschwale, Reaux, & Čech, 1996).

Kinds of Experience

If the sheer amount of experience on a series of tests may or may not alter underlying comprehension, what about the *kinds* of experience that a chimpanzee might have? For example, what if one were to grow up an ape, but in the company and culture of humans? Could this experience significantly alter the epigenetic course and outcome of a chimpanzee's cognitive development? Tomasello turns our attention to "enculturated" apes as a case in point and wonders whether more experience with humans might not have significantly affected our apes' performances. Some of his own research on imitation suggests that chimpanzees raised with humans and adorned with their material culture develop at least a rudimentary understanding of the intentions of others (Tomasello, Savage-Rumbaugh, & Kruger, 1993).

To begin, I do not possess any strong a priori feelings about whether apes reared by humans undergo profound psychological reorganizations. However, I do not think that the strength of the empirical evidence to date forces us to take the view that an understanding of intentions—which on Tomasello's reading normally eludes the chimpanzee's mental system—can be imbued by raising them with another species whose cognitive/behavioral system is so constructed. The best data sets available with which to evaluate his hypothesis concern the capacity for imitation, which his studies suggest may not be present in wild chimpanzees but may emerge in chimpanzees enculturated with humans (Tomasello et al., 1993). However, there is at least some data that indicate that imitation may be present in chimpanzees no more enculturated (and some certainly less so) than our subjects (see Custance, Whiten, & Bard, in press; Whiten, Custance, Gómez, Teixidor, & Bard, 1996). Thus, it may be difficult to argue that enculturation is responsible for the *de novo* appearance of true imitative abilities.

In other domains such as chimpanzees' understanding of seeing, the evidence that Tomasello marshals in support of the enculturation hypothesis is often circumstantial, limited to quasi-experimental approaches that have been only incompletely reported (i.e., Premack's, 1988, account of how one of four chimpanzees attempted to remove an experimenter's blindfold before enlisting his assistance on a task). In addition, because the relevant developmental inputs are not precisely specified (although for Tomasello they are related to pragmatics more that language per se), it is difficult to know how to apply the enculturation hypothesis uniformly. For instance,

Tomasello rejects Premack and Woodruff's (1978b) landmark study of a highly enculturated adult chimpanzee's understanding of intention by arguing that the task (selecting a photograph that fulfills the goal of an actor on videotape struggling to solve a staged problem) could be completed by simple associative rules. However, he does not mention Premack and Woodruff's (1978a) theoretical refutation of this objection or Premack and Dasser's (1991) empirical data with young children demonstrating the implausibility of this counterinterpretation. My point is neither that Premack and Woodruff's (1978b) tests with Sarah definitely established that chimpanzees understand intention nor that Tomasello's enculturation hypothesis is incorrect. Rather, it should be clear that accounts of this kind are a double-edged sword for the enculturation hypothesis—if Sarah merely used clever association-based algorithms to solve this task, then a similar kind of analysis could be applied to other enculturated apes on other tasks as well.

Having expressed my reservations about the enculturation hypothesis, let me now explore some of the implications of accepting (or rejecting) it. First, if mere exposure to the conventions and material culture of humans can transform their understanding of others so profoundly, then we might no longer wish to entertain the idea of dedicated cognitive modules that evolved explicitly for the purpose of generating inferences about the mental states of others (e.g., Baron-Cohen, 1995; Fodor, 1992; Leslie, 1994). If chimpanzees (and other nonhuman primates) normally do not possess the cognitive abilities that modularity theorists envision as being controlled by specialized, evolved brain modules, but if enculturated chimpanzees do, then obviously the underlying neural systems subserving theory of mind in humans did not originally evolve for that purpose. If true, our neurobiological account of theory of mind would need to be recast in terms of the evolution of developmental systems that retain a high degree of plasticity in the face of different epigenetic environments.

From the opposite point of view, however, let us make the assumption that the typical absence of the expression of theory of mind in chimpanzees can be generalized to older animals and furthermore reflects the unique evolution in humans of some kind of neural system controlling its expression. That is, let us imagine that cognitive specializations in theory of mind (regardless of its exact neural basis) evolved in humans after our lineage split from the line leading to the African apes (Povinelli & Preuss, 1995). How, then, could we account for the apparent cultivation of theory of mind in enculturated apes? Again, one possibility is that its presence is more apparent than real. Although the behavioral end product may be similar, enculturated apes may have arrived via a different route. One of the clear lessons from this *Monograph* is that with sufficient experience (perhaps of exactly the type received by enculturated animals) chimpanzees may act as if they understood the intentionality of seeing, for example, without really

doing so. In this case, the "curious limitations" that Tomasello notes in the enculturated ape's social understanding may not be limitations at all. Rather, they may merely reflect the holes left in the nonmentalistic cognitive scaffolds constructed by one species forced to cope with a different species that expects such social understanding.

Finally, a limitation of the enculturation hypothesis is that it never clearly specifies which inputs are critical or how these inputs inbue the ape's cognitive system with an understanding of the mental life of others. This is not a criticism unique to the enculturation hypothesis. Rather, it exposes broad, conceptual uncertainties surrounding our ideas of how theory of mind abilities are constructed and develop during the course of human ontogeny. This aside, consider Tomasello's closing comments that the evolution of human culture may be dependent on the evolution of skills related to social cognition. I agree, and we have presented a detailed argument about how these cognitive skills explain both the workings and evolution of universal aspects human ethics and morality and the absence of such systems in chimpanzees (Povinelli & Godfrey, 1993). Yet this agreement masks some important unresolved issues concerning both the nature of the adaptations that produced these systems in humans in the first place and the epigenetic environments that trigger their development now. One possibility is that this epigenetic system reflects underlying changes in the genes controlling development of specific neural systems that were selected for during the course of human evolution. But if this is true, how can mere enculturation with humans trigger its expression in a species that descended from ancestors who (by definition) split off from the line leading to humans prior to the evolution of this adaptation? One possible answer is that forcing the ape's developing system to accommodate the content and practice of certain aspects of human culture may somewhat alter its typical ontogenetic course. The theoretical challenge for the enculturation hypothesis is, How much and how fundamentally can this course be altered?

What about Language?

Thus far, the causal role of human language has played a small part in this discussion of the enculturation hypothesis. Yet it is quite possible that some aspects of training apes in regimes designed to have them produce and comprehend the semantic and syntactic features of human language may result in deep, foundational changes in their conceptual systems. For example, Premack (1988) argued that training apes to use a symbol for same/different judgments about objects produced a cognitive system that could understand the relations critical to solving analogies. Is it similarly possible that exposure to certain aspects of a human language system (pragmatic,

semantic, or syntactic) may directly alter the ape's theory of mind? This question raises a variant of the even broader classic question concerning the relation between language and mind: Is theory of mind possible in the absence of a communication system with the semantic and syntactic features of human language?

Assessing the effect of language training on a chimpanzee's theory of mind raises a number of extremely difficult challenges. First, the strong version of the language transformation hypothesis hinges on whether an ape's exposure to and use of aspects of a human language system forces it to comprehend the various aspects of that language system (pragmatic, syntactic, semantic) in the way humans do. Second, there is the question of whether mental states (as concepts) are or can be represented in the absence of a linguistic code that can compress as much information as human language. This, of course, is the classic mind-language problem. It is possible that apes (and other nonhuman animals) have mental codes that allow for the use of images to represent something like epistemic and nonepistemic mental states. Indeed, we could even imagine that apes possess a syntactic device for parsing and then productively reordering such images in order to generate novel strings of images (ideas). (On the other hand, I have often wondered whether many of the differences between apes and humans are the consequence of a very limited ability of apes to engage in such syntactic construction and deconstruction of the behavior stream. That is, perhaps humans have specialized in an ability to decompose a given segment of the behavior stream in almost infinitely unique ways. Certainly, the sophistication of this ability would have direct implications for the extent to which the behavior of apes is exclusively governed by a complex interaction of social dispositions and learned procedural rules.)

Yet even if wild chimpanzees possess a system for manipulating an image-based code—a system powerful enough to produce a sufficient distance of sign from referent that we would wish to gloss some given image as a representation of a "mental state"—we still must acknowledge the fact that if language-trained apes come to use symbols to compress information on the order of magnitude that humans do, then perhaps they may find themselves in a position to distance the observed actions of others from idealized instances of those actions. In short, if language-trained apes come to use symbols as a shorthand for compressing many ideas about objects or events into a single unit, this could provide a similar vehicle for generating a kind of shorthand, heuristic code for representing the potential actions of others—a code perhaps not so different from the foundational aspects of our own folk psychology. If this were the case, then the limitations on their ability to understand mental states would be related to the extent to which their inherited representational systems can accommodate compression of action potentials into "ideas" about such unobservable states. Fur-

thermore, it would not be exposure to the pragmatics of human language that would facilitate an understanding of mental states but rather the semantic and syntactic features of language. Of course, all this hinges on our ability to resolve the first point raised above. That is, if the behavioral productions of these abnormally reared, language-trained apes are the result of learned rules and procedures unrelated to the syntactic and semantic features of human language, we would have no reason to entertain these ideas (for empirical data bearing on this question, see Savage-Rumbaugh et al., 1993).

Are Apes "Mere Behaviorists"?

In this Reply, I have continued to maintain our position that our apes' behavior on our tests is best predicted by a learning framework. However, as is his gift, Hobson grapples with the details of this claim and exposes what is, in truth, a far more complicated portrait than we have outlined. He is correct in assuming from our disclaimers that we are in full agreement that a spectrum of possibilities exists concerning young chimpanzees' and young humans' understanding of the subjective aspects of any given facet of mental life. One's understanding of the social world is not necessarily either behaviorist or mentalistic. Elements of both can and do coexist alongside each other in humans, and they may in other species as well.

In addition, to be a mentalist one may have either a quite elaborate understanding of mental states and events (as do adult humans) or but a fledgling appreciation of the subjective aspect of self and other (as perhaps do human infants and children). Similarly, Tomasello underscores our point that the apes' understanding of seeing need not be as sophisticated as young 3-year-olds' in order to qualify as genuinely mentalistic (after all, even our 2½-year-olds did not always do well). Indeed, as we noted in our second alternative hypothesis, there may be ways in which apes (or human infants and children) can understand attention other than understanding the eyes or face as portals through which attention emanates. Furthermore, as Hobson points out, chimpanzees or human infants may also react to others in ways that suggest that they register (or directly perceive) the subjectivity of another's expressive attitude or feeling, without being able to represent that this is to what they are responding. Indeed, in posing these challenges, Hobson has anticipated that much of our current work (despite our earlier protestations of the theoretical difficulties involved) bears directly on both these questions (Povinelli et al., 1996; Povinelli & Eddy, 1996, in press).

There are several aspects of Hobson's insistence that we consider all points on the "spectrum of possibilities" that I embrace wholeheartedly.

First, as he points out, we need an account of how infants develop an appreciation of the subjectivity of self, "selves," and the mental world in general that is not discontinuous—an explanation that does not leave the infant suddenly leaping from behaviorist to mentalist. Important in this kind of explanation will be the kinds of psychological configurations that an infant's developing brain can sustain short of a full appreciation of the content and character of adult mental states, attitudes, and events. Although in places I have trouble understanding why some of his "attitudes" could not be described in strictly procedural terms, in general the point is well taken. Second, I agree that we need an explanation of how and when evolution has acted to produce this kind of ontogenetic system in the first place. Indeed, John Cant and I have recently attempted to provide an evolutionary account of at least one strand of this puzzle, the evolutionary emergence of understanding the self as a causal agent (Povinelli & Cant, 1995).

The qualitative versus quantitative change problem that Hobson grapples with is an old one in comparative psychology. Indeed, it is so old that many comparative psychologists dismiss it as archaic, a nonissue. I disagree. As we pointed out in Chapter I, this problem remains a conceptual challenge to developing a comparative psychology of mind. Yet in expressing my appreciation of the difficulty of these issues I must also express my concern about how one might interpret Hobson's comments. One line of reasoning typically goes like this: (1) humans possess a subjective understanding of self and other; (2) these abilities evolved (they were not divinely given); (3) chimpanzees are demonstrably very closely related to humans; and, therefore, (4) chimpanzees must possess at least rudiments of the same qualitative kind of subjective understanding of self and other that humans develop. As I have pointed out elsewhere, this inference does not follow (see Povinelli, 1993). When it comes to understanding the continuous chain of processes that have led to the evolution of a given biological system or structure, it is one thing to speak of precursors, foundations, or building blocks, but quite another to tackle the more difficult problem of function. Because we inhabit only a very narrow slice of evolutionary time, we are privy to only a thin cross section of the diversity that it has produced—a cross section that, because of extinction and genetic/morphological/behavioral evolution within populations, does not retain all the continuous branches of the tree we wish to reconstruct (see Fig. R2). Extinction, coupled with the imperfections of the fossil record, virtually guarantees that during periods of rapid evolutionary change we can never accurately identify the exact origin and subsequent elaboration of a given trait. Simply put, when it comes to evolution, discontinuous patterns do not necessarily imply discontinuous processes.

The point I wish to make most strongly is this: chimpanzees and other great apes and nonhuman primates may or may not share with humans

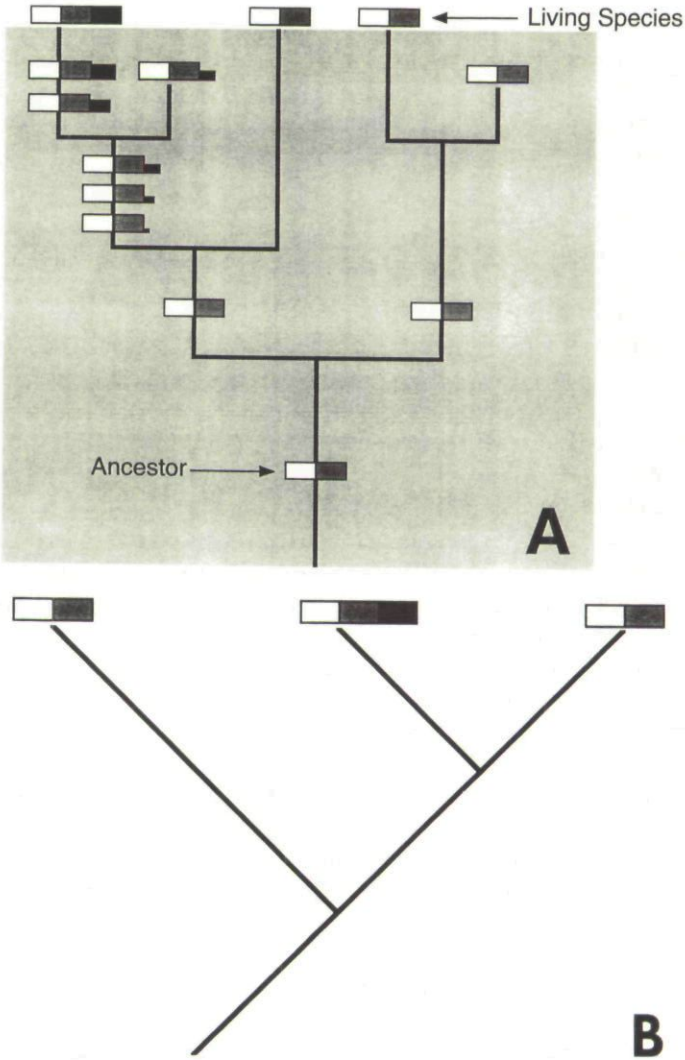

FIG. R2.—In *a*, a modern radiation of some arbitrary clade of species is shown, complete with all ancestors and descendants. In addition, the evolutionary emergence of a hypothetical cognitive-behavioral system (dark box) is tracked through evolutionary time from its initial appearance to its later elaboration. For purposes of clarity, this elaboration is depicted as the trait increasing in size. Extinct species inhabit the shaded region, whereas the three living species appear in the unshaded area. In *b*, just the living representatives of the radiation are shown as a part of a modern clade. Note that in *a* there is no dramatic step from the absence of the new system to its presence (although even here its initial appearance may have occurred as a result of a qualitative shift in function from previous, related systems). However, in *b* the appearance of the novel system appears discontinuous because of the fact that living species almost never comprise true ancestor-descendant relationships and because of the imperfections of the fossil record. Indeed, in the case of the evolution of cognitive systems, it is not clear whether even a pristine fossil record could yield sufficient morphological information (i.e., brain endocasts) to infer the presence or absence of the cognitive system.

limited, perhaps foundational aspects of theory of mind. However, our preconceived ideas about the rate and timing of evolutionary change cannot produce an a priori answer to the question. In this respect, evolutionary reconstructions are part of a historical science that is impenetrable to intuitions about how things ought to have evolved.

An Alternative View

In the final section of this Reply, I wish briefly to sketch an alternative view to the fairly simplistic notion of the evolution of cognitive development that has permeated my own thinking and research (as well as, I believe, that of others) for a number of years. I hope to show how the confusing array of similarities and differences that we see between humans and chimpanzees may be understood within the framework of modern biological ideas about the evolution of development (i.e., heterochrony).

Imagine either that humans have exclusively evolved a specialization in theory of mind or that the ancestors of modern humans inherited a fairly limited mentalistic psychology that was then radically elaborated and transformed during the course of their separate 5–7-million-year radiation as a separate lineage of bipedal apes. In either of these cases, humans can be envisioned as possessing an elaborated psychological system for generating, sustaining, and revising representations of the intentions, goals, desires, beliefs, thoughts, and activities of themselves and others. If this is true, what are we to make of spontaneously occurring behaviors (such as desception) that humans and apes share in common but that humans alone appear to understand in a thoroughly mentalistic manner? Take the case of gaze following (see Experiment 12, Chap. IV; Povinelli & Eddy, 1996). Both humans and chimpanzees have been shown to be capable of elaborate forms of this ability. Some researchers see the emergence of this behavior as signaling the arrival of the capacity for joint or shared attention in human infants (e.g., Baron-Cohen, 1995). However, one skeptical possibility is that this behavior has nothing whatsoever to do with the infant's understanding of the attentional focus of others. Rather, it could be that a later-developing psychological system in humans (their theory of mind) allows for a retrospective interpretation of gaze in this manner—a reinterpretation that is never allowed for by the ape's psychological system.

However, it is possible that we need to be willing to think differently about certain behaviors depending on whether it is an ape or a human infant that is displaying them. It could be that when apes engage in so-called shared attention behaviors, they do so without a co-occurring assignment of meaning, but that when human infants engage in the very same behaviors, such an assignment is made. Before I am criticized for using a double

standard, it would be worthwhile to entertain the possibility that gaze following, deception, and the like are fairly ancient *behavioral* mechanisms that evolved long before anything like theory of mind had evolved. Maybe it is our species alone that has evolved a capacity to represent each other's mental states and in doing so has been placed in a position of being able to reinterpret behavioral patterns that evolved long before we did. Consider the following. After following someone else's gaze, when is it that we wonder what it is that she is looking at? Is the gaze-following response triggered by an inference about the mental state, or does exactly the opposite causal relation hold? Or perhaps it is something intermediate between the two. We have argued that much of the utility of theory of mind occurs in exactly this retrospective manner—as a mechanism for planning future actions, not as a means of solving immediate social problems (Povinelli & Povinelli, 1996). In short, where some researchers tend to describe social primates as having evolved a coherent theory of mind to cope with on-line social problems, we envision nearly the opposite. What social mammals need most is to act quickly, and what better for this than the ability to form a powerful set of automatized, procedural social rules from which they can quickly select an appropriate series of behavioral scripts. In this view, theory of mind is merely a reflective, after-the-fact, added bonus.

Thus, in our species (at least) a connection appears to exist between expressions of certain behaviors and a retrospective mentalistic interpretation of those behaviors. This may reflect the operation of cognitive innovations peculiar to our species that in the course of human ontogeny now emerge alongside more ancestral behaviors that evolved in isolation from this system. There are really at least two possible versions of this idea. First, the innovations in cognitive development may express themselves very early in the developing human infant, so that during each phase of cognitive development a psychological mechanism related to theory of mind allows for a qualitatively different kind of interpretation of behavior than what is present in other species (Povinelli, Zebouni, & Prince, 1996). A second possibility is that the behaviors that some researchers like Hobson interpret as evidence of a kind of intersubjectivity or shared affect or attention are really no more than behavioral algorithms that emerge through epigenetic interactions during the course of infant development. As opposed to tinkering with the earliest phases of cognitive development, selection may have modified later phases to gradually produce an epigenetic system that could support representations of mental states. For example, it is quite compatible with a skeptical view of the developmental literature in this area that human infants must await their second year of life before they develop the evolutionarily novel cognitive structures that allow them to see their own behavior and that of others in genuinely mentalistic terms.

Although I believe that it is a hypothesis worthy of consideration, I

cannot demonstrate with any degree of certainty that the alternative view that I have sketched here is correct. Neither can Tomasello demonstrate unequivocally that enculturated apes develop an interpretive framework that shares many of the features of our own theory of mind system. Likewise, Hobson cannot establish without doubt that 1-year-old human infants are engaged in a genuinely mentalistic attitude toward others or that chimpanzees share this stance. Nor, indeed, do I think any of us would want to make such strong claims in the face of so young a science. Rather, their Commentaries and this Reply follow a more rewarding (and ultimately more productive) path of using current results to understand the shape of the remaining theoretical possibilities.

References

Amsterdam, B. (1972). Mirror self-image reactions before age two. *Developmental Psychobiology,* **5,** 297–305.

Asendorpf, J. B., & Baudonniere, P.-M. (1993). Self-awareness and other-awareness: Mirror self-recognition and synchronic imitation among unfamiliar peers. *Developmental Psychology,* **29,** 88–95.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism.* Cambridge, MA: MIT Press.

Bischof-Köhler, D. (1988). Uber der Zusammenhang von Empathie und der Fahigkeit, sich im Spiegel zu erkennen [On the association between empathy and ability to recognize oneself in the mirror]. *Schhweizerische Zeitschrift für Psychologie,* **47,** 147–159.

Bischof-Köhler, D. (1994). Selbstobjektivierung und fremdbezogene emotionen [Self-objectification and other-oriented emotions]. *Zeitschrift für Psychologie,* **202,** 349–377.

Custance, D. M., Whiten, A., & Bard, K. A. (in press). Can young chimpanzees (*Pan troglodytes*) imitate arbitrary actions? Hayes and Hayes (1952) revisited. *Behaviour.*

Fodor, J. A. (1992). A theory of the child's theory of mind. *Cognition,* **44,** 283–296.

Gallup, G. G., Jr., & Suarez, S. D. (1986). Self-awareness and the emergence of mind in humans and other primates. In J. Suls & A. G. Greenwald (Eds.), *Psychological perspectives on the self* (Vol. **3**). Hillsdale, NJ: Erlbaum.

Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences,* **16,** 1–14.

Gopnik, A., & Meltzoff, A. N. (1986). Relations between semantic and cognitive development in the one-word stage: The specificity hypothesis. *Child Development,* **57,** 1040–1053.

Gopnik, A., & Meltzoff, A. N. (in press). *Words, thoughts and theories.* Cambridge, MA: MIT Press.

Johnson, D. B. (1982). Altruistic behavior and the development of the self in infants. *Merrill-Palmer Quarterly,* **28,** 379–388.

Leslie, A. M. (1994). ToMM, ToBy, and agency: Core architecture and domain specificity in cognition and culture. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture.* New York: Cambridge University Press.

Lewis, M., & Brooks-Gunn, J. (1979). *Social cognition and the acquisition of self.* New York: Plenum.

Lewis, M., Sullivan, M. W., Stanger, C., & Weiss, M. (1989). Self-development and self-conscious emotions. *Child Development,* **60,** 146–156.

Povinelli, D. J. (1993). Reconstructing the evolution of mind. *American Psychologist,* **48,** 493–509.

Povinelli, D. J. (1994). How to create self-recognizing gorillas (but don't try it on macaques). In S. Parker, R. Mitchell, & M. Boccia (Eds.), *Self-awareness in animals and humans.* Cambridge: Cambridge University Press.

Povinelli, D. J. (1996). Chimpanzee theory of mind? The long road to strong inference. In P. Carruthers & P. Smith (Eds.), *Theories of theories of mind.* Cambridge: Cambridge University Press.

Povinelli, D. J., Bierschwale, D. T., Reaux, J. E., & Čech, C. G. (1996). *Do adolescent chimpanzees understand attention as a mental state?* Manuscript submitted for publication.

Povinelli, D. J., & Cant, J. G. H. (1995). Arboreal clambering and the evolution of self-conception. *Quarterly Review of Biology,* **70,** 393–421.

Povinelli, D. J., & Eddy, T. J. (1996). Chimpanzees: Joint visual attention. *Psychological Science,* **7,** 129–135.

Povinelli, D. J., & Eddy, T. J. (in press). Factors influencing young chimpanzees' recognition of "attention." *Journal of Comparative Psychology.*

Povinelli, D. J., & Godfrey, L. R. (1993). The chimpanzee's mind: How noble in reason? How absent of ethics? In M. Nitecki & D. Nitecki (Eds.), *Evolutionary ethics.* Albany: State University of New York Press.

Povinelli, D. J., & Povinelli, T. J. (1996). Review of *Mindblindness: An essay on autism and theory of mind,* by Simon Baron-Cohen. *Trends in Neuroscience,* **19,** 299–300.

Povinelli, D. J., & Preuss, T. M. (1995). Theory of mind: Evolutionary history of a cognitive specialization. *Trends in Neuroscience,* **18,** 418–424.

Povinelli, D. J., Zebouni, M. C., & Prince, C. G. (1996). Ontogeny, evolution, and folk psychology. *Behavioral and Brain Sciences,* **19,** 137–138.

Premack, D. (1988). Minds with and without language. In L. Weiskrantz (Ed.), *Thought without language.* Oxford: Clarendon.

Premack, D., & Dasser, V. (1991). Perceptual origins and conceptual evidence for theory of mind in apes and children. In A. Whiten (Ed.), *Natural theories of mind.* Oxford: Blackwell.

Premack, D., & Woodruff, G. (1978a). Author's response. *Behavioral and Brain Sciences,* **1,** 616–629.

Premack, D., & Woodruff, G. (1978b). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences,* **1,** 515–526.

Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., & Rumbaugh, D. M. (1993). Language comprehension in ape and child. *Monographs of the Society for Research in Child Development,* **58**(3–4, Serial No. 233).

Tomasello, M., Savage-Rumbaugh, E. S., & Kruger, A. C. (1993). Imitative learning of actions on objects by children, chimpanzees, and enculturated chimpanzees. *Child Development,* **64,** 1688–1705.

Whiten, A., Custance, D. M., Gómez, J. C., Teixidor, P., & Bard, K. A. (1996). Imitative learning of artificial fruit processing in children (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology,* **110,** 3–14.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.