The Self in Infancy: Theory and Research
P. Rochat (Editor)

CHAPTER 9

The Unduplicated Self

DANIEL J. POVINELLI

University of Southwestern Louisiana
New Iberia Research Center

"The present...is really a part of the past — a recent past — delusively given as being a time that intervenes between the past and the future. Let it be named the specious present, and let the past...be known as the obvious past."
– E.R. Clay

If pressed, developmental psychologists offer 18–24 months as the period in which humans infants develop a self-concept, citing this as the age at which infants are able to recognize themselves in mirrors (Amsterdam, 1972; Lewis & Brooks-Gunn, 1979; for other achievements suggesting this as the onset of self-conception, see Kagan, 1981). Although contributors to this volume may not completely disagree with this view, most are united in the belief that long before 18 months, human infants process and store many kinds of information about themselves — a belief that highlights the need for understanding the preconceptual origins of the self in early infancy. In their view, the kinds of self-conceptual behaviors that 18- to 24-month-olds display (such as self-recognition in mirrors) are seen as the expression of an already sophisticated self-representational system.

In this chapter, I shall steer a narrow path between these two caricatures of the emergence of the self-concept. On the one hand, I will acknowledge the sophistication of the 18- to 24-month-old's self-knowledge by contrasting it with preconceptual forms of knowledge present much earlier in infancy. Indeed, I will defend the proposition that by 18–24 months, most children have developed a qualitatively new system for encoding information about the self. On the other hand, I will emphasize the immaturity of the 2-year-old's self-knowledge by offering the hypothesis that a striking difference exists between the 18- to 24-month-old's self-conceptual system and that of older preschool children. I outline a theoretical model which argues that once the young child's conception of self is in place at 18–24 months, it continues to develop from an initial system of self-

representations that is largely restricted to the here and now (the "present self") to a later system that is capable of temporal differentiation (the "proper self"). I propose that these two hypothetical self-constructs emerge asynchronously during the course of development. It is the latter construct that allows the child to knit together historical instances of him- or herself into a unique, unduplicated self. It is what, in William James' (1890/1950) words, allows our consciousness to say, "*I am the same self that I was yesterday*" (p. 332; italics in original).

Throughout this essay, I will provide empirical grounding for the theoretical discussion by examining the reactions of human infants and children, as well as chimpanzees and other nonhuman primates, to visual presentations (and representations) of themselves (e.g., seeing themselves in mirrors). I take this approach not because I believe that mirrors are the only methodological tool for understanding the development of the self — they most certainly are not. Instead, I focus on mirrors because the different kinds of reactions that live versus delayed visual feedback of the self provoke from human infants and children of various ages, as well as other species, provide an important window into the architecture of their knowledge about the self. In addition, the self-recognition paradigm lends itself easily to our comparative program, which seeks to identify homologous cognitive developmental pathways among the great apes and humans. Furthermore, I shall maintain that self-recognition in mirrors reflects a much deeper understanding of the self than simply knowing what one looks like.

Representational Development

The foundation of my theoretical model of the development of the self-concept is that its emergence at 18–24 months is but one manifestation of a more general capacity for representation. In other words, although the self-concept is an important domain of inquiry for developmental psychologists, its development is by definition not possible until the capacity for representation proper emerges. In order to sustain such a view, I need to provide a framework for thinking about the development of representation in human infancy. The formal account of transformations in the infant's representational abilities that I adopt is derived largely from the theoretical work of Olson & Campbell (1993), although the particular application of this to self-representation is my own.

Representations versus Schemata

During infancy, humans develop the capacity to construct action schemata. Schemata can be thought of as neural control programs that are triggered by stimuli in the external world. These programs control motor output, and hence

they warrant a description as causal structures. Throughout the first 18 months of life, the human infant advances from the use of relatively simple action schemata (reaching, grasping, head-turning) to the later elaboration and complex deployment of such schemata (one hand reaches to open a box, the other reaches inside the box, grasps an object, pulls it toward the mouth). Gradually, such schemata become automatized and can be deployed readily and in rapid succession in the presence of the relevant stimuli. According to sophisticated rules of generalization, these schemata can be triggered in appropriate "novel" contexts. The progressive construction of these behaviors requires an increasingly sophisticated distinction between self and the environment, presumably mediated by proprioceptive feedback. But despite their eventual sophistication, these internal neural structures are not representations of the external world:

> Schemata are activated by a present object or event; they are causally connected to that object or event, but they do not represent that object or event. Their activity is tied to stimulating conditions; consequently, there is no need for the activating condition to be held in mind. It is present, and therefore present to the mind (Olson & Campbell, 1993, p. 15).

Thus, despite the complexity of the behavior that these early schemata can execute, they are not representations proper — they are not held in mind in the absence of the objects of perception that trigger them in such a way that they can be *intrinsically connected* (see below) to other schemata. Olson and Campbell (1993) conclude that it is not until 18–24 months that mental representation — the ability to create an intentional connection between a schema (in the absence of its stimulating condition) and another object or event in perception — becomes possible. On this view, earlier emerging demonstrations of object permanence using habituation-dishabituation techniques (e.g., Baillargeon, 1987) would not be interpreted as evidence of an active ability to create an intentional relation between two schemata, although it may be evidence of early implicit connections among schemata (see Olson, 1993). Although I recognize this point to be debatable, I shall stipulate this implicit/explicit distinction without further justifying it.

A wide range of behaviors blossom in the child at this age, including symbolic play, mirror self-recognition, simple acts of altruism, self-descriptive utterances, mastery smiles, statements about desires, sophisticated gaze-following, linguistic markers of the failure to self-generated plans, acts of intentional cooperation, development of the self-conscious emotions, an explosion in naming skills, second-order classification skills, and an understanding of referential focus, just to name a few (Leslie, 1987; Amsterdam, 1972; Lewis & Brooks-Gunn, 1979; Zahn-Waxler & Radke-Yarrow, 1982; Kagan, 1981; Bartsch & Wellman, 1995; Bischof-Köhler, 1988; Butterworth & Jarrett, 1991; Gopnik, 1982; Brownell & Carriger, 1990; Lewis, Sullivan, Stanger, & Weiss, 1989;

Macnamara, 1982; Langer, 1986; Baldwin, 1993). In contrast to schemata, mental representations are a means of "maintaining a relation [an intentional connection] with an object or event *in its absence*" (Olson & Campbell, 1993, p. 14; italics in original). (Note that this meaning of representation captures its linguistic origins as the "re-presentation" of an object or event in its absence.) One defining feature of the capacity for representation is the ability to actively create relations between things directly perceived and things only conceived. Such relations are called propositions. Propositions can embody either linguistic or imaginal relations between separate schemata. Thus, with the advent of representational ability, the infant now has the capacity to construct many types of propositions. For example, Olson (1993) speculates that the explosion in naming skills displayed by 18- to 24-month-olds is the direct result of their new-found ability to form propositions, expressed as instance or category relations: "This [object of perception] *is a* ball [held in mind]." Similarly, classification becomes possible on the basis of hierarchical categories: "This ball [object of perception] *is a* toy [held in mind]." Other relations become possible as well: "The ball [held in mind] *is under* that cup [object of perception]." Also, and critical to my later hypothesis, causal attribution should become possible: "This event [object of perception] *was caused by* that action [held in mind]."

Representing Propositions

Olson (1993) has argued that there is a further elaboration of the child's capacity for representation between 3 and 5 years (see also Olson & Campbell, 1993). He argues that, whereas by 18–24 months infants have the resources to hold in mind one schema (independent of the environmental context that stimulates it) while their perceptual system attends to something else and creates an intentional relation between the two, by 4 years of age children develop the ability to represent the propositions themselves. One consequence of this is that it allows propositions to be evaluated (true or false) in relation to perceived states of affairs in the world. In this view, these representations are not limited to mental states, but in fact are part of a broader developmental transition. Olson (1993) has used this model to explain how once a young child has the conceptual resources to hold in mind more than one representation at a time, he or she is able to pass most standard theory-of-mind tasks. Indeed, I should be careful in committing to a particular version of how the child becomes capable of embedding representations within representations. In recent years, a number of proposals have been put forward to account for changes during the preschool years in young children's ability to cope with representations (Flavell, 1988; Forguson & Gopnik, 1988; Perner, 1991; Lillard, 1993). As a case in point, consider Flavell's (1988)

interpretation of his research on the development of young children's understanding of the appearance-reality distinction:

> Children of this age [3-year-olds] also believe, as we generally do, that each object or event in the world has only one nature — one "way it is" — at any given point in time. It cannot be two or more very different, mutually contradictory, and incompatible things at the same time; rather, it can only be one thing. Consequently, it makes no sense to them to hear something described as being radically different than the single way it "is" (with "is" not differentiated from "seems to them at the moment") (Flavell, 1988, p. 245).

Olson's (1993) theoretical position offers one account of the underlying reason for this inability of children younger than about 3 or 4 years of age to cope with such situations. But regardless of the particular view to which one subscribes, once the child achieves the functional ability to understand that individual things can have multiple states, profound changes quickly follow in their understanding of objects, events, and mental states. I shall argue that this transition has an equally profound impact on the child's conception of self as an entity with a personal history and future.

The Present Self

Based on these foundations, let us assume that from birth forward (the period covered by most contributors to this volume) the developing infant elaborates upon the construction and deployment of schemata. By 18–24 months, however, the infant has the capacity to hold in mind a representation of the self (a self-concept) while its perceptual system is directly engaged with objects or events in the world. This initial system continuously updates and replaces its self-representations, but because it cannot hold in mind two representations at the same time, it is unable to store former representations of the self *in relation to* these new ones. Borrowing from William James (1890/1950), I shall refer to this representation of the self as the "present self." According to my model, the most primitive manifestation of this present self is the child's ability to form mental declaratives that amount to self-descriptions of their physical or mental states (especially such as agency and desire) (e.g., "I am building a house," "I am hungry," "I cause(d) this," "I want this"; Kagan, 1982; see Bartsch & Wellman, 1995, for extended analysis of these kinds of spontaneous utterances in 2-year-old children). For now, I define this representation of the self's immediate state as

(1) S_i ,

where S denotes representational-based knowledge of the self and the subscript i fixes the location at the immediate location in time. This initial self-conceptual system is familiar to adults (James, 1890/1950). Indeed, part of the current hypothesis is that this present self is not replaced by a different conceptual system later, but rather remains in place throughout later developmental elaborations (see below).

Despite this similarity to our adult understanding of self, my assumption that simultaneous comparisons of different representations are impossible implies that self-representations are largely "on-line"; what is represented is only a single representation of the child's physical and mental states. Thus, by definition this system carries with it no capacity to integrate previous mental or physical states with current ones. This is not to say that previous self-representations play no role here; some subset of them may well be stored in memory and may even provide default inputs into the child's S_i . Thus, as the child's current perceptual information about itself changes, its self-representation is updated to match those changes. Some of these successive self-representations may be relatively stable, or at least may experience strong continuity from one to the next, and hence they may not be updated very much or very often. On the other hand, some of these self-representations (such as the child's desires) may be updated relatively frequently. This allows us to formally define S_i as a temporally localized, integrated set of self-experiences represented by the child:

(2) $S_i = (\{s_i, \text{physical}\}, \{s_i, \text{psychological}\})$

Several important clarifications about this view of the initial self-concept are needed. First, the present is not a "knife-edge," but instead carries with it some temporal degradation (James, 1890/1950). In other words, our conception of the here and now can never truly be the here and now because by the time we direct our attention to that instant, it has already vanished. This is what E.R. Clay (cited by James, 1890) meant by his reference to the "specious present." Thus, any description of the present self must recognize a time corridor into the immediate past and into the immediate future. It is difficult at this point to specify the exact time dimensions along which the child (or the adult) carves up the present, but ultimately it will be critical to do so. Indeed, in a later portion of this essay I offer some suggestions as to how we might experimentally investigate this issue.

The second clarification concerns the fact that memories of the relatively distant past clearly exist in very young infants, and children who are 2 to 3 years of age have verbal access to such memories and can recall details of events that occurred at remote points in the past (see Nelson, 1989, 1991, 1993). Thus, this model must grant the 18- to 24-month-old child access to at least some of those

memories. Any attempt to define a present self must explain how children can have access to these past representations if S_i is temporally truncated in the manner I have stipulated. I speculate that from the 2- to 3-year-old's perspective, these stored self-representations are "atemporal" in the sense that they have no temporal or causal relation to S_i . Thus, many (but not all) of these previous states can be readily recalled (e.g., Nelson, 1989; Gopnik & Slaughter, 1991), but they need have no relation to S_i . I will explore the implications of this issue later, but for now I acknowledge this access and incorporate it into the present self. Finally, by 18–24 months of age, young children have also developed the capacity for imagination and can also talk about events likely to happen in the future. Again, however, I suspect that these imagined states are atemporal in the same sense as representations of previous states.

The Proper Self

Next, I consider how the emergence of the capacity to hold in mind several representations simultaneously (which develops in the later preschool years) may interact with the on-line self-conceptual system described above. Just as the advent of the shift from schemata-based to representation-based knowledge allows for the emergence of the present self, I hypothesize that an additional capacity to embed propositions within representations has an important impact upon the child's self-conceptual system. Such domain-general changes, which allow children to simultaneously compare multiple representations, ought to have important ramifications for the child's self-representations. Using Olson's (1993) terminology, I speculate that with the emergence of the ability to hold in mind more than one representation, the child's representational system begins to organize what were successive on-line self-representations (S_i) as a separate concept, which I define as the proper self:

(3) S_p

I speculate that this higher-order representation of the self serves as an organizing concept for all past states of the self (defined as S_{i-n}) and all future imagined states of the self (defined as S_{i+n}), and links them together into an organized temporal progression, where n denotes a unit of time. This allows us to expand (3) more formally:

(4) $S_p = (\{S_{i-n}, n > 0\}, \{S_i\}, \{S_{i+n}, n > 0\})$

This formalism implies several things. First, each previous state of the self that is stored in memory (S_{i-n}) is a member of a set that in totality comprises a more general representation of the self (S_p). It also implies that, as n approaches 0, these states more closely approximate the organism's current state. Although there are obvious exceptions, an organism with a folk understanding of causality could (as a heuristic) assume that, for any previous state of themselves that they consider, as n approaches 0 that state has an increasing similarity to S_i.[1] Similar reasoning applies to imagined future states.

To summarize, I speculate that the ability to simultaneously hold in mind previous representations of S_i allows for the child to establish a causal relation between these former states and S_i. In effect, this causal relation establishes the irreversible arrow of time as part of the child's folk psychology. S_p can thus be understood as the higher-order representation of the self, which holds the self-concept together as an enduring entity through time with a past and a future. This view is consistent with some recent views concerning the timing of the onset of autobiographical memory (see Nelson, 1993). One important consequence of this is the construction of a temporal corridor along which the self progresses from past to present and (via imagination) into the future. It may be that this time line is viewed as deterministically moving in one direction, or it may be viewed as being cyclical in nature. At least, however, when the child considers or confronts a former or future representation of him- or herself, the child can group it as a particular instance of S_p. Thus, it becomes possible for the child to do more than simply recall and verbally report previous physical and mental states from the past, but to understand how one's current state its causally determined by one's previous states. Likewise, and more generally, the child can simultaneously consider multiple physical and/or psychological states of the self as referring to the same concept or entity, which I have defined as S_p. One probable consequence of this is that *coincident with the emergence of the capacity to embed representations of the self within propositions about the self, the child discovers its own ontogeny.* [2] Or, to borrow from William James (1890/1950) yet again, it creates the capacity to conceive of an "unbrokenness in the stream of selves" (p. 335).[3]

Self-recognition in Comparative Perspective

Having established the foundations for three different ways in which the developing infant or child processes and stores information about the self (one preconceptual, two conceptual), I now wish to move on to begin to consider the following question: Why is it that self-recognition in mirrors only occurs in organisms that possess conceptual knowledge of the self? After answering this, I will attempt to

show how the formalisms established above can be used to generate some new (and counterintuitive) predictions about limitations of young preschoolers' ability for self-recognition. I begin by providing a quick sketch of what is currently known about self-recognition in mirrors in human infants and nonhuman primates.

Self-recognition in Nonhuman Primates

In a widely known series of experiments, Gallup (1970) exposed chimpanzees and several different species of macaques to mirrors for a period of 10 days. Initially, most of the animals responded by engaging in a number of species-typical social reactions, such as threatening, lip-smacking, play behavior, and sexual presentations. However, after 2 or 3 days, some striking species differences began to emerge. Unlike the macaques, chimpanzees began to engage in a number of behaviors that Gallup labeled as "self-directed." These behaviors suggested that the subjects had discovered that the real source of the image was themselves. For example, the animals were reported to engage in repetitive movements of the limbs and exaggerated facial movements, and to use their hands to explore parts of themselves that they had never seen before (teeth, nose, ano-genital region) — all while carefully monitoring the mirror image (see Figure 1). After 10 days of mirror exposure, the subjects were anesthetized and marked on the upper eyebrow ridge and ear with a bright red dye, which offered little or no olfactory or tactile cues. The significance of Gallup's procedure was that when the animals recovered they would have no way of knowing that they were so marked.

After complete recovery from the anesthesia, the subjects were observed for a 30-minute control period in the absence of a mirror. Any attempts to touch the marked regions were noted. Next, the mirror was introduced, and again the number of mark-directed contacts was recorded. The chimpanzees made few if any contacts to the marked areas during the control period, but made a number of contacts in the mirror test. Indeed, the subjects often inspected their fingers immediately after making contact with the marks, despite the fact that the marks left no olfactory or tactile cues. This pattern of results supported Gallup's initial impressions that the chimpanzees had correctly discovered the source of the mirror image. In contrast, the monkeys who were marked and tested in the same fashion made no attempts to touch the marks. These basic findings have been replicated many times, and have been extended to include orangutans (Gallup, McClure, Hill, & Bundy, 1971; Lethmate & Dücker, 1973; Suarez & Gallup, 1981; Calhoun & Thompson, 1988; Lin, Bard, & Anderson, 1992; Povinelli, Rulf, Landau, & Bierschwale, 1993). Likewise, the failure to find self-recognition in members of primate species outside the great ape-human clade has been widely replicated, despite some ingenious attempts to make the source of the image more obvious (Benhar, Carlton, & Samuel, 1975; Gallup, 1977a; Gallup, Wallnau, & Suarez, 1980; Anderson, 1984;

Itakura, 1987a, 1987b; Anderson & Roeder, 1989; Marchal & Anderson, 1993). Indeed, although many reasonable methodological criticisms have been raised concerning the repeated negative findings of self-recognition in monkeys, most have been empirically addressed and have been found wanting (see Gallup, 1977a; Gallup & Suarez, 1986; Gallup, Wallnau, & Suarez, 1980; Anderson & Roeder,

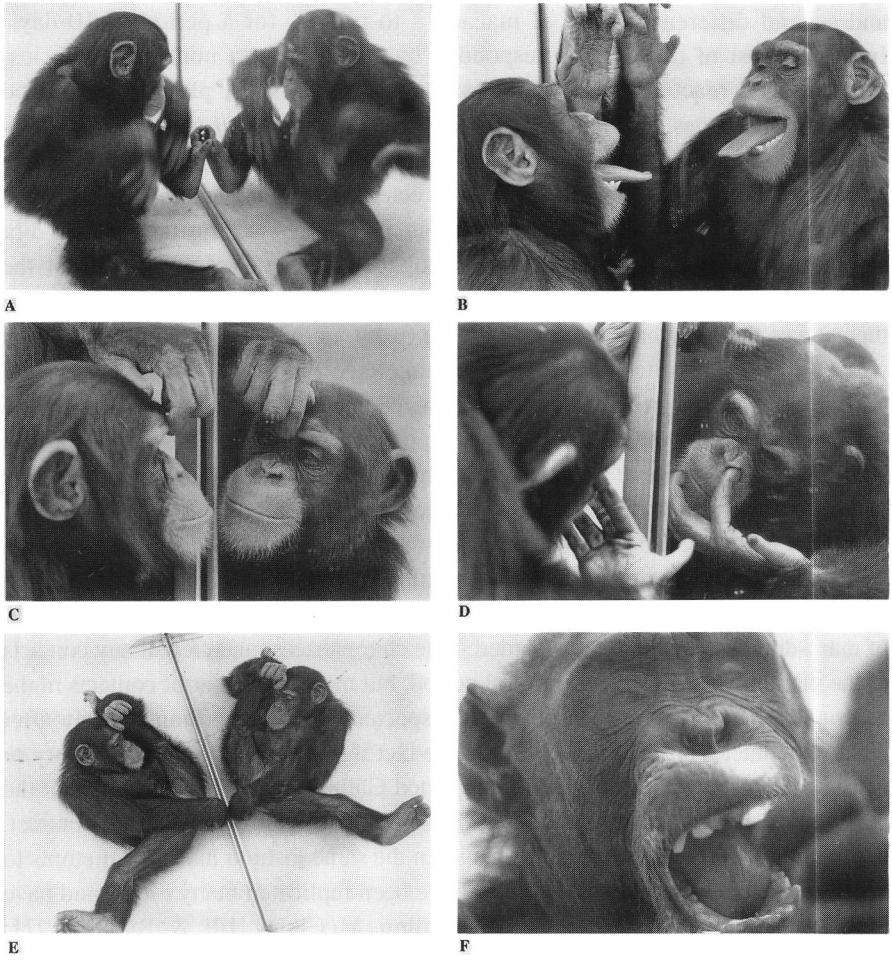

FIGURE 1. Chimpanzees are capable of using mirrors to engage in (a) contingent body movements, (b) contingent facial movements, and (c-f) self-exploratory behaviors. Some of these behaviors may indicate that the animals understand that the image in the mirror is equivalent to themselves (see text for details). Photographs by Donna T. Bierschwale.

1989; Itakura, 1987a, b; Anderson, 1986; Povinelli, 1989). To be sure, there are debates about the distribution, ontogeny, patterns of emergence, and underlying cause of self-recognition in chimpanzees (Swartz & Evans, 1991; Lin, Bard, & Anderson, 1992; Povinelli et al., 1993; Mitchell, 1993). However, I do not believe we can seriously question the basic finding that many chimpanzees are capable of using a mirror to explore parts of themselves and are able to pass well-controlled mark tests (see Gallup et al., in press).

Although Gallup (1970) originally suggested a conglomerate of self-directed behaviors that were indicative of self-recognition, researchers have recently distinguished between contingent body, contingent facial, and self-exploratory behaviors (Lin et al., 1992; Povinelli et al., 1993). In particular, Povinelli et al. have argued that contingent body and facial movements are not good indicators of self-recognition; in other words, they do not predict the presence of behaviors in which the animals seem to be using the mirror to explore parts of themselves (self-exploratory behaviors), nor are they good predictors of passing a mark test. Conversely, not all behaviors that might appear "self-exploratory" necessarily indicate that the subjects are using the mirror to explore themselves. Povinelli et al. (1993) suggest that some of this activity may be ambient-level self-grooming or scratching, or indeed some might be heightened levels of scratching caused by arousal (a phenomenon common in nonhuman primates, e.g., Maestripieri, Schino, Aureli, & Troisi, 1992). New techniques have been devised to control for this problem by recording the reactions of chimpanzees to mirrors or live video feedback, versus prerecorded videotape stimuli of other chimpanzees in a similar setting. These techniques reveal that young chimpanzees that display contingent body and facial movements (but not self-exploratory behavior) to the mirrors will initially display these "contingent" behaviors to the prerecorded videotape of other chimpanzees (Eddy, Gallup, & Povinelli, 1995). I shall return to the significance of this finding later, but for now I will simply note that we interpret this as evidence that long before chimpanzees are capable of recognizing themselves in mirrors, they learn through proprioceptive and kinesthetic feedback that they can control the movement of the mirror image. Thus, the young animals appear to learn procedural rules for manipulating the behavior of the "other" animal they see in a mirror, and when confronted with the same situation except that the stimuli is a video of others, they are duped into executing the same procedural rules.

Development of Self-recognition in Humans

Amsterdam (1972) independently invented the mark test for use with human infants. Although similar to Gallup's, her methodology differs in several critical ways (see Gallup, 1994). Given that most subsequent investigators have followed a variant of her methodology, what follows is a description of a typical test of

infant self-recognition (e.g., Amsterdam, 1972; Schulman & Kaplowitz, 1977; Bertenthal & Fischer, 1978; Lewis & Brooks-Gunn, 1979; Zazzo, 1982; Johnson, 1983; Asendorpf & Baudonnière, 1993). Infants are marked (usually by their mother) by having a location of their face (nose, forehead, chin, or cheek) wiped with a cloth containing some kind of wet marking substance (usually a cosmetic). Next, they are immediately presented with a mirror. The typical finding is that very few children younger than about 16–18 months display evidence of self-recognition. However, by 18–24 months approximately 60–70% of infants will "pass" the test. What constitutes passing can range from reaching up to touch the mark, providing a correct verbal label for the image ("me" or the child's proper name), or even drawing attention to the mark in the mirror. Gallup (1994) has noted that not all of these are valid measures of self-recognition. The formal model I outline in this chapter will highlight some of the conceptual problems with accepting some of these behaviors as evidence of self-recognition. Distinguishing among the divergent kinds of measures of self-recognition in human infants is important because not all indicate the presence of S_i. Although I shall return to this issue later, as prelude to my model it is important to note that only one of these measures (reaching up to touch the mark) can be construed as clear evidence that the infants understand that the image in the mirror refers to (or is about) themselves.

Self-recognition and Self-conception

The studies of self-recognition in children and nonhuman primates described above have traditionally been used as a mean of exploring an organism's self-concept. For instance, Gallup (1975) concluded that self-recognition in mirrors "would seem to necessitate an already established identity on the part of the organism making that inference" (p. 330).[4] Working with young children, Lewis and Brooks-Gunn (1979) argued that mirror self-recognition is indicative of the presence of a form of objective self-awareness.

Mitchell (1993) has recently criticized the terminology used by Gallup and others and has proposed two models to account for the phenomenon. Both have attempted to explain mirror self-recognition in terms of mental operations in which the child's self-concept per se plays little or no role. It is important to note that Gallup has made two separate claims. His initial claim was that self-recognition in mirrors was only possible with a well-integrated self-concept (Gallup, 1970, 1975, 1977b). The exact nature of this self-concept was not specified. More recently, Gallup (1982) argued that if self-recognition in chimpanzees implicates some kind of self-concept, and if that self-concept is elaborate enough, then chimpanzees may be capable of using their own experiences to model the experiences of others. He therefore developed a model predicting that

chimpanzees should engage in mental state attribution, but that other species which fail to show evidence of self-recognition should not. Although a large portion of our own research agenda is designed to investigate whether chimpanzees attribute mental states to others, this issue is not considered in this chapter (see Povinelli & Eddy, in press). Instead, I focus on: a) the kind of self-knowledge that makes self-recognition in mirrors possible; and b) how that self-knowledge is distinct from that of older children.

Self-recognition in Mirrors Revisited

I now use the formalisms adopted earlier to explain why self-recognition in mirrors is restricted to organisms that have developed the general capacity for representation (and in particular, the capacity to represent the self).

Self-recognition in Mirrors Derives from Conceptual Knowledge of the Self

Let us begin with organisms that meet the following criteria: they a) are old enough to possess the capacity for representation (approximately 18–24 months); and b) have not yet learned to recognize themselves. First, like younger infants, the 18- to 24-month-old's perceptual system detects the contingency between its motion and the motion in the mirror.[5] However, assuming that the infants have already applied their representational capacities to the self (that is, they have constructed S_i), I propose that the infant will perform an agency mapping in the form of a causal proposition very rapidly: "That action by the mirror-image (object of perception) *was caused by* me (S_i)." Thus, because they have the capacity to hold in mind a representation of S_i while perceptually attending to other things, their mirror-image is tagged by the child as connected to (or being about) S_i. I speculate that very rapidly the infant takes the additional step that not only is the image about S_i, but it is equivalent to S_i. This equivalence is constructed precisely because every feature (physical and/or psychological) that the infant can represent about the self is also true of its mirror-image. For example, the 18-month-old's limited attributional capacities (e.g., the ability to attribute desire), causes it to attribute to the mirror-image psychological states equivalent to its own. As the child reaches for a ball in front of a mirror while automatically monitoring (on-line) its state of desire for the ball, so too does it attribute to the mirror-image (which is reaching for the same object) the same desire-goal relation. The child therefore constructs the equivalence:

(6) mirror-image $\equiv S_i$

Thus, with the construction of a self-concept of the type outlined earlier, the answer to the question, "What is causing that?" becomes clear: "I (defined as S_i) am." This is a conceptual form of knowledge of the self's agency.

Finally, what about the infant prior to 18–24 months? Although younger infants are sensitive to contingency in the manner outlined by Bahrick and Watson (1985) and Meltzoff (1990), the absence of a capacity for mental representation leaves them with no ability to infer a referent for the mirror-image. If we assume that these infants only have access to schemata-, physiological-, and proprioceptive-based information about the self, then by definition they have no resources that allow them to perceptually attend to one object or event (for example, their image in a mirror) while simultaneously relating that thing to an object or event not present (see above). Thus, when they encounter an image of the self in a mirror, they may learn any number of procedural rules that result in specific payoffs. When infants (or animals) first look into mirrors and see objects or events that are behind them in real space, they may be duped into responding as if those objects or events were where they appear in the mirror. However, they can easily learn procedural rules whereby they respond to the appearance of reflected objects by turning around to their real location in space. For example, numerous researchers have demonstrated that with sufficient experience with mirrors, human infants and animals who do not pass a mark test can use a mirror to direct their visual or manual searches to the real location of the rewarding object or event (e.g., Anderson, 1986; Itakura, 1987a, 1987b; Povinelli, 1989; Robinson et al., 1990). Likewise, they can also learn that when they move, so does their mirror-image. Indeed, because of this contingency, infants too young for self-recognition may even learn (through proprioceptive feedback) that they can control the movement of their mirror-image (see above; Povinelli et al., 1993; Eddy et al., 1995). Yet despite the construction and deployment of such sophisticated schemata, the image has no relation to anything else. More to the point, for the organism observing itself in a mirror, the image has no relation to the self precisely because (by definition) the self (S_i) cannot yet be represented.

To summarize, I am hypothesizing that mirror self-recognition occurs because organisms with general representational abilities also form robust representations of themselves. This allows them to understand that the image in the mirror moving with them is about them, and ultimately refers to or is equivalent to them. Note, however, that I am not claiming they understand that their mirror image is a representation of themselves. Indeed, according to this model, organisms do not understand this until considerably later.

Recognizing Parts of the Self?

Some researchers will argue that I have sidestepped the question of how other phenomena such as facial or bodily recognition, imitation, or more importantly an understanding of mirror correspondence, contributes to self-recognition in mirrors. After all, some authors have identified some or all of these as key factors in providing a coherent account of self-recognition (Guillaume, 1926/1971; Parker, 1991; Mitchell, 1993). Yet according to the model outlined here, these phenomena are incidental to mirror self-recognition, not necessary features for a system to display behaviors indicative of self-recognition: self-exploratory behaviors in front of mirrors and/or reaching up to touch a previously unknown mark on the self.

First, consider the question of how the child comes to know that the facial features seen in the mirror belong to him or her. Some researchers see this as the whole task of self-recognition in mirrors. But I believe that careful analysis reveals that this is a trivial component of the question of mirror self-recognition. On the one hand, infants too young to construct an S_i still may be able to correctly label the featural cues they see in a mirror by using the words *me* or their proper name. As many authors have pointed out, verbal labeling of this kind may merely mean that the child or infant has learned from its parents that the face (that set of features) is correctly labeled using their proper name (i.e., "Mary") or the first person pronoun *me* (Gallup, 1975; Bigelow, 1981; Anderson, 1984b). In addition, by as early as 5 months of age infants may also show other evidence of discrimination of their own face from strangers, presumably because through previous exposure to mirrors their facial features are simply more familiar than a stranger's (e.g., Fadil, Moss, & Bahrick, 1993). On the other hand, the same infant who is now old enough to form the equivalence relation in (6) but who has not yet done so may have no idea what his or her facial features look like — this, despite the fact that the infant has been able to *discriminate* his or her face from others for months.

Some will insist that I have still not explained how children can know that the facial features seen in the mirror belong to them before they recognize themselves (Mitchell, 1993). Of course, it is true that prior to the very first time a child recognizes herself in a mirror certain aspects of her physical self-representation may be incomplete. But I do not see a need to specify a unique inferential or deductive process by which the child incorporates details of his or her facial appearance (for example) into an updated $s_{i,\ \text{physical}}$. For the prototypical case, I will assume that the mirror-naive organism has little or no default inputs about its own facial features. The model that I am advancing argues that coincident with the advent of representational abilities, the child creates a cohesive self-identity that includes aspects of her physical and mental states. The child is constantly updating these self-representations, just as she does with other objects

and events in the world. Thus, prior to the child understanding that the image in the mirror is equivalent to herself, she has a concept of self (S_i) and hence her encounter with a mirror (which is marked with perfect contingency between what the child sees in the mirror and S_i) leads the child to form the predicate relation (6) between S_i and the mirror-image. Whereas earlier the child's perceptual system detected the contingency between her actions and external events (including those in a mirror, on a live video monitor, or those simply "caused" by its direct actions upon the world), by 18–24 months that perceptual system has been articulated to a conceptual system, which includes self-representations. Thus, children are able to hold in mind on-line internal representations of self (S_i) as subject, with their perceptual systems free to attend to the mirror-image as the predicate. Once such a proposition is established, new information (in this case knowledge about the details of one's facial features) is simply incorporated into S_i as the representation is updated.

Thus, the process through which the child updates his $s_{i,\ \text{physical}}$ from information provided via the mirror image would not seem to differ from a child updating his representations of other things in the world. For instance, imagine a 2-year-old child whose only experience with automobiles was from looking out the front window of his or her house as cars drove up into the driveway. With sufficient experience, the child could clearly form the concept of car and sufficiently generalize it to novel exemplars of the class. But now imagine that the child has the opportunity to go outside and actually explore a car for the first time. Consider all of the details that the child will now discover about each car that is visited (they have seats, license plates in the back, etc.). The child's new concept of automobiles includes new features that he or she did not know of before. In my view, this is no different than what occurs when children learn the details of their own facial features. Once the tagging occurs as described in (6), I suspect that children simply update their representation of their physical appearance.

Up to this point I have provided an account of how an organism that is learning to recognize itself for the first time reaches the equivalence relation specified in (6). What about its subsequent interactions with mirrors? Here a slightly more complicated situation arises. Now that the organism has a more complete representation of its physical appearance, the exact triggering cause of the equivalence relation (6) is difficult to specify. Depending on the deployment of the attentional resources of the infant, the focus may be exclusively on physical similarity, or it may be on the psychological similarity of agency, which is embedded in both contingency and desire reflected in equivalent actions on the environment. However, as I shall show later, placing these two dimensions of the self in conflict with one another in the context of a visual re-presentation of the

self has the potential to reveal the underlying differences among children of different ages.

Must Organisms that Display Self-recognition Understand Mirrors?

If the model I am offering is correct, then it means that one of the most widespread ideas about self-recognition — that it requires some practical knowledge of how mirrors work — is incorrect. Mitchell (1993), for instance, argues in both of his models of mirror self-recognition that an organism needs to understand the property of "mirror correspondence," which is defined as the knowledge that "mirrors reflect accurate and contingent images of objects in front of them" (p. 298). He sees mirror correspondence as one of the necessary conditions for the organism to infer or deduct (depending on which of his theories one examines) that the image in the mirror is itself. However, according to the model offered here, the construction of the equivalence outlined in (6) does not require any such specific competence. In order for the child to construct an equivalence between her mirror image and her existing self-representations, she need not understand that mirrors accurately and contingently reflect what is in front of them. This would explain why there does not appear to be a reliable correlation between understanding mirror correspondence (an infant's ability to use mirrors to localize events or objects out of their perceptual field) and passing the mark test (Zazzo, 1982; Loveland, 1986; Robinson, Connell, McKenzie, & Day, 1990). Some infants who test positive for self-recognition using a mark test appear to understand this property of mirrors (or at least have formed procedural rules that make it appear as if they understand this property); others do not. Further, it would also explain why children reared in cultures that have few or no fabricated mirrors seem to be capable of recognizing themselves (passing a mark test) after only a few minutes of their first exposure to a mirror (Priel & de Schonen, 1986). My interpretation of these data is that children who pass the mark test may or may not understand the affordances of mirrors because understanding mirrors is not necessary in order to arrive at the equivalence relation given in (6).

Another reason why the child's interactions with mirrors have seemed mysterious is because the mark test appears to involve finding something that is hidden from view (i.e., the mark on the face). For instance, Bertenthal and Fischer (1978) speculated that the correlation between the development of object permanence and self-recognition they obtained was probably because both involve skills related to "the ability to search for hidden objects" (p. 49). Mitchell's (1993) body-part objectification, object-permanence theory is perhaps the most formal statement of this view. But an implication of the formalism adopted in this essay is that the difference between a 14-month-old child who does not pass the mark test and an 18-month-old who does, has nothing to do with respective unsuccessful

versus successful "searches" for the mark. Further, as noted above, it is not that the 18-month-old has figured out that mirrors accurately reflect what is in front of them (an understanding of "mirror correspondence"), but rather that they have an existing on-line S_i and can form the proposition described in (6). But none of this implies that they have actively searched for, and then found, the real location of what they are seeing in the mirror. Rather, they are simply using the mirror to gain access to either previously unknown aspects of the self (as in the case of the onset of self-recognition) or to explore some alteration of a previous self-representation of the face (as in the case of the mark test after self-recognition has occurred).

Thus, children (or chimpanzees) who recognize themselves in mirrors need not understand all (or even most) of the reflective properties of mirrors. They need not understand (although they may) that mirrors reflect things that are in front of them, nor need they understand mirrors as representational devices. They simply need the conceptual capacity to form the predicate relation that the mirror image is equivalent to S_i. This alone induces the child with a self-concept to explore his face, not the surface of the mirror. Another way of looking at this would be to say that the child could have a quite stable S_i but could be quite confused about how or why it is framed in glass. On the other hand, children (or apes or monkeys) who have sufficient experience with mirrors early in life may come to form sets of procedural rules (schemata) for responding to mirrors. Conversely, this account also explains how both young infants and monkeys without representational capacities of the type discussed here could, with sufficient experience, learn procedural rules to react to mirrors appropriately but still not display behaviors indicative of self-recognition (e.g., Itakura, 1987a, 1987b; Robinson et al., 1990). Thus, self-recognition in mirrors requires neither an understanding of the reflective properties of mirrors in general nor the understanding that the mirror image is a *representation* of S_i in particular.

Finally, let me note that this model shares some features with Mitchell's (1993), in that proprioceptive matching contributes to the formation of the equivalence relation (6). But after that the models depart. Mitchell assumes that the limiting factor for the onset of self-recognition in young infants is either: a) the absence of an elaborate enough "kinesthetic-visual matching" capacity (which is present at birth, and quite elaborated by 9–14 months, e.g., Meltzoff & Moore, 1977; Meltzoff, 1990); or b) the absence of an understanding of mirror correspondence. In direct contrast, the model offered here specifies that: a) an understanding of mirror correspondence is not necessary to form the equivalence relation (6); and b) the kinds of sensitivity to contingency that are necessary for self-recognition are in place long before 18–24 months. What has not yet developed, and thus what can be described as the limiting factors for self-

recognition, are the representational abilities necessary for the construction of the self-concept (S_i).

Predictions of the Theory and Preliminary Tests

If the theory outlined above is to be seriously entertained, it should generate some previously unarticulated predictions about the behavior of young children and chimpanzees at various stages in their development. Below I explore some of these predictions, one of which involves a previously unpredicted asynchrony in children's capacity to recognize themselves using live versus delayed visual feedback.

Interactions with Mirrors Before the Emergence of the Present Self

First, the theory predicts that there ought to be little correlation between an organism's ability to use mirrors to locate objects or events using mirror cues and its ability to pass a mark test or display mirror-mediated self-exploratory behaviors. Thus, before they are able to pass a mark test of self-recognition, human infants and chimpanzees given sufficient experience with mirrors should be able to learn to use mirrors to locate objects that (for example) are really behind them. Likewise, other species in which self-recognition has never been demonstrated should also be able to use mirrors in this fashion. As I have indicated above, both of these predictions seem to be borne out by the existing data. As a rather large case in point, take our previous research with elephants: Although they displayed no evidence of recognizing themselves in mirrors, they showed a very sophisticated ability to use mirrors to locate hidden objects (Povinelli, 1989).

A second prediction that we have explored empirically is the idea that organisms within a species whose members are capable of self-recognition in mirrors ought to be able to learn to manipulate a mirror-image using information about the self that is available though preconceptual channels (proprioceptive feedback). Thus, schemata-based self-knowledge should allow an organism to form procedural rules about its actions and contingent consequences in another location. Consistent with this view are the results from the studies described earlier, which indicate that before chimpanzees pass a mark test or display self-exploratory behaviors in front of mirrors, they learn to manipulate the image by engaging in contingent body and facial movements (see Figure 1; Povinelli et al., 1993). That their interpretation is not based upon a equivalence relation of the type outlined in (6) is apparent from the fact that they attempt the same manipulations when the stimuli are prerecorded images of other chimpanzees. Thus, the procedural nature of the self-knowledge is exposed.

Self-recognition in Mirrors and Beyond

Next, a provisional acceptance of the model can generate some surprising predictions about what might happen if we manipulate either a) the contingency and/or b) the featural similarity of the child's self-image. Let me begin with contingency. Recall that 2- to 3-year-olds are assumed to possess a representational capacity such that they are able to represent their physical states, personal agency, desires, and perceptual experiences (and possibly their knowledge states) as S_i. Second, their sensitivity to contingency leads them to construct the equivalence proposition that $S_i \equiv$ mirror image. But consider what would happen to such a system's response to a "re-presentation" of the self that was not contingent with its current actions (or desires, or knowledge states, for that matter). As a case in point, let us imagine that the image is of a series of events just previously performed by the child, described as the set:

(7) $(\{S_i - 1\}, \{S_i - 2\}, \ldots \{S_i - n\})$

In this case, the detection of physical similarities of the image (i.e., bodily or facial recognition) should lead the system to provide a verbal description of the image using either *me* or the child's proper name. Indeed, if its attentional resources were maximally devoted to the featural cues, the organism might even momentarily form the equivalence relation in (6). However, as soon as the child's attentional resources focused upon the other (and I assume more salient) aspects of the re-presentation such as its agency, he or she should conclude that the image is not equivalent to S_i. After all, neither agency nor desire-goal states appear to match; therefore, there is no reason to establish a straightforward proposition linking the two. An obvious alternative route to connecting the two is available to adults: relating the object of perception (the events on the monitor) to the representation S_i through the representation S_p. Yet the model stipulates that this route is not available to the 2- to 3-year-old. Nonetheless, these younger children should still be capable of identifying the images using their proper name (or even *me*, if the child treats *me* and his/her proper name as denoting that set of featural cues).

How might we go about determining if young children actually reason in such a dissociated fashion about the re-presentations of a previous image of themselves? There are several ways, but let me first address a paradigm that we have explored (Povinelli, Landau, & Perilloux, in press). First, imagine that a child is playing a simple game where he or she looks under cups for stickers. One experimenter is directly playing with the child, and the other is sitting next to the child, praising her or him and patting him or her on the head after each sticker is found. Imagine also that the procedure is being videotaped so that the child's head is clearly

visible. On the third trial, the second experimenter uses the act of praising the child as the vehicle to covertly place a large sticker on top of the child's head. Several additional trials are conducted to ensure that the child does not detect the marking. Next, the child is invited to watch what he or she did on television. Thus, two minutes after the child was marked, the child sees the events replayed on the monitor.[6]

The critical question is, of course, what do the children do when the playback reaches the point at which the experimenter places the sticker on their heads? Contrary to what one might expect, the theory predicts that between 18–24 months and 3 years or so, children should not reach up to remove the sticker. They should be too young to construct S_p, and thus should be unable to infer that the sticker they see on their forehead in the video is currently on their head; thus, for them, S_i is not causally related to the set $(\{S_{i-1}\}, \{S_{i-2}\}, \ldots \{S_{i-n}\})$. Yet the theory predicts that if these same children are placed in front of a mirror, they should remove the sticker almost immediately precisely because they can form the proposition, $S_i \equiv$ mirror image. Recall, however, that the theory also makes the explicit prediction that these younger children should have no trouble whatsoever identifying the child on the prerecorded videotape by using *me* or their proper name.

Older 3-year-olds and most 4-year-olds ought to be in very different position. Their ability to consider multiple representations of the same thing simultaneously should allow them not merely to label the image using *me* or their proper name, but also to infer that what they are witnessing on the videotape is a particular instance of their proper self, a former S_i temporally adjacent with their current S_i. Thus, an inference of what is true of any given S_{i-n} may also be true of S_i is likely to be drawn, especially (as we have seen) as n approaches 0. The theory therefore predicts that unlike their younger counterparts, they ought to reach up to search for the sticker after the tape reveals the experimenter placing it on their heads.[7]

Preliminary Tests

To date, we have conducted three tests of the idea just described (Povinelli, Landau, & Perilloux, in press). In the first experiment, we executed the procedures exactly as described above using forty-two children ranging from 2 to 4 years of age as the subjects (ten 2-year-olds, sixteen 3-year-olds, sixteen 4-year-olds). As the model predicted, none of the 2-year-olds and only 25% of the 3-year-olds reached up to take the sticker off their heads after the tape showed the experimenter placing it on their heads. In contrast, a full 75% of the 4-year-olds reached up to remove the sticker within an average of 7 seconds after viewing that part of the tape. These results are especially dramatic when they are contrasted with studies of

mirror self-recognition, where up to 80% of the 2- and young 3-year-olds would be expected to pass the test.

The second experiment was identical to the one described above, except that instead of filming the children using videotape, we took Polaroid snapshots of the children at two junctures in the procedure. The first photo was taken as the experimenter was praising the child and covertly placing the sticker on her head. The second photo was taken after the control trials. One of the experimenters introduced the child to a large stuffed gorilla that she had never seen before and explained that the other experimenter was going to take another picture, this time of the child, the gorilla, and the experimenter together. Thus, two snapshots were available to show the children: one that clearly depicted the experimenter placing the sticker on the child's head, the other depicting the child (with sticker on head) sitting with the gorilla and the experimenter. We hoped that these images would force in a more direct way the relation between what had just happened and the current state of affairs. Approximately 2–3 minutes after the second picture was taken, the main experimenter showed the child the photographs one at a time. During the presentation, the experimenter asked her a series of standardized questions designed to determine her ability to identify the images correctly. Finally, if the subject had not reached up to remove the sticker by the end of the presentation of the second photograph, she was given the stuffed gorilla again, presented with a mirror, and invited to look at themselves along with the main experimenter. A total of 60 children participated in this study, with 15 children in each of 4 age groups: young and old 3-year-olds, and young and old 4-year-olds.

Consistent with the first experiment, only 13% of the young 3-year-olds reached up to their heads to remove the sticker while the photographs were presented and the questions were asked. In contrast (and in full accord with the theory), 85% of the young 3-year-olds who did not reach up while looking at the photographs did so when they were presented with the mirror. In contrast, by 4 years of age, 80–90% of the subjects reached up while looking at the first photograph.

Other aspects of the results also support some of the predictions of the theory. For example, even the youngest children were able to provide a "correct" verbal label for their image: When the experimenter pointed to their image and asked, "Who is that?," 75% of the children in the youngest age group responded by stating their name or using the personal pronoun *me*. However, there was an intriguing significant developmental difference in the use of the personal pronoun *me* versus their proper names. The youngest children appeared not to discriminate between using their proper name or the personal pronoun *me*. In contrast, by 4 years of age, the response profile had shifted completely. Virtually all of them responded by using the personal pronoun *me*. Further linguistic evidence that the

younger children did not interpret their photographic image as relating to S_i comes from an analysis of the kinds of possessive verbal descriptions the children used when the experimenter pointed to the sticker in the photograph and asked, "Where is the sticker right now?" Most of the children in the youngest age group (after having just "correctly" identified the image by stating *me* or by using their proper name) described the sticker as being on *his* or *her* or *the* head, whereas only a few described its location using the first person possessive pronoun (*my head*). In striking contrast, not a single 4-year-old child used the third-person possessive pronouns. These patterns of answers are very consistent with the view that the younger children recognized and had a verbal label for their featural cues, but did not relate the image to their current present self. In addition, they show that even when the youngest children's attention was explicitly focused on the sticker in the image, they still failed to reach up to remove it.

The third study produced results that suggest that the model as presented is incomplete. In this study, we directly compared two groups. One received delayed video feedback of the self after the child was marked with a sticker as in the first experiment, and the other received live video feedback of the self after the child was marked in an identical fashion. Thus, both groups observed themselves for two minutes with the stickers on their heads, with the only difference being the contingency of the image. We tested 48 subjects ranging in age from 2–3 years (mean age for both group = 35 months), with 24 subjects in each group.

Although the group differences were in the direction predicted, the difference was not statistically significant (62% of the children in the Live Feedback condition reached up to remove the sticker, and 37% of the children in the Delayed Feedback condition did so, $p < .07$). The results surprised us for two reasons. First, a higher percentage of 2- and young 3-year-olds passed the Delayed Feedback test than had in our previous studies, and far fewer passed in the Live Feedback condition than would be expected based on previous studies of mirror self-recognition. However, there are several intriguing possibilities that could explain these findings in the context of the theory. First, it is possible that the greater percentage of children passing the test in the Delayed Condition than did previously could be due to the fact that they did not see their image transformed from one state (not marked) to another state (marked) as they did in the first experiment involving delayed video. Ironically, we had originally speculated that seeing the whole marking event might make it easier for the younger children; in retrospect, the theory seems to predict the opposite. First, recall that the theory stipulates that the younger children can only consider one representation of themselves at a time. However, as they watch themselves on the monitor without a sticker, it is still true that the featural dimensions of the image are consistent with their current default inputs regarding $s_{i,\,\text{physical}}$. Thus, if all of their

attentional resources were focused on the featural cues at the expense of the contingency/agency cues, they could momentarily form the proposition that the image is equivalent to S_i , but would have no reason to reach up to their heads because there are no marks yet. However, as soon as the child sees the experimenter reach up and act upon the image, he concludes that it cannot be about him because no one is putting a sticker on his head. Notice, however, that in the last experiment the child in the delayed group only saw himself in one state — with a sticker on his head. However, in this case, when his attentional resources are initially devoted to the featural cues, this might cause him to form the equivalence relation, and thus reach up and remove the sticker. But as soon as his attention is drawn again to the agency discrepancy, he should be just as strong in his conviction that the image is not about S_i .

Although we are just beginning the experiments that will be necessary to tease these issues apart, it is worth reflecting upon the continuing sorts of spontaneous verbal comments made by these younger children when confronted with their delayed images. For example, one young girl in the Delayed Feedback test reached up immediately to her head when her image appeared, but then asked in confusion several seconds later why the girl on the screen did not take the sticker off her head, too. The general point is that the theory does not exclude the children from shifting back and forth between an interpretation of equivalence and nonequivalence, depending upon whether they focus (or the experimental paradigm forces them to focus) on the featural or contingency cues. Given the consistent negative pattern that has emerged with children 3 and younger confronting noncontingent stimuli, I conclude that the contingency factor is the most salient.

Finally, why did so few children reach up in the Live Feedback condition? We speculate that it is because the best kind of stimuli to cause the equivalence relation is live symmetrical (or specular) feedback (i.e., mirrors, or live video in which the normal reversed image is made mirrorlike). We have been testing the idea that part of the motivation for the equivalence relation concerns not sensitivity to contingency, but a certain form of identical contingency — symmetrical contingency. This predicts that if the Live Feedback had been of a different kind — a specular kind — many more children would pass the test. Although there are somewhat uninteresting reasons why this might be the case (i.e., children develop scripts for dealing with mirrors, and antispecular images depart from those scripts), there may be more fundamental explanations. Although there is not room to expand upon this idea here, this sensitivity may be present in early infancy and may reflect natural selection for organisms' ability to detect when others are behaviorally connected or linked to them.

A final issue I wish to raise concerns the temporal breadth of the present self as conceived by the 2-year-old child. To cast the question in empirical terms,

imagine that we repeated our original delayed test of self-recognition, but instead of having the age of the child as our independent variable, we manipulated the interval between the time at which the marking event actually happened and the moment when the child saw the playback of the events. How close to perfect temporal contingency would be necessary before 2-year-olds consistently formed the proposition relating the image and S_i ? Such experiments (if properly designed) might provide some insight into the duration of the troubling "specious present" — at least for the 2-year-old child. Although we have yet to begin such studies, I suspect that this window is very narrow indeed.

Self-conception in Evolutionary Perspective

The theory outlined in this chapter has implications far beyond the development of the self in human infancy. It also sets the stage for answering questions concerning the nature of the self-concept in other species that provide some initial reason to suspect that they form some kind of self-concept. Chimpanzees and orangutans represent obvious choices, given that members of both species have been demonstrated to show every bit as compelling evidence of self-recognition as 18- to 24-month-old human infants.

In this essay, I have committed to a domain-general view of the development of representational capacities in human infancy and childhood. However, ignoring the exact rate of development, it is not completely clear if these synchronies in development exist in other species, such as chimpanzees. Cast in slightly different terms, it is unclear which cognitive-developmental pathways are dissociable. Elsewhere I have tried to assess the current evidence concerning the homologous aspects of representational development between humans and chimpanzees as reflected in the domain of theory of mind (Povinelli, in press). A fair reading of the research to date leads to the conclusion that it is still too early to determine which aspects of cognitive development typical of the 18- to 24-month-old human infant occur in chimpanzees. Even for the ones that we can be reasonably certain exist in chimpanzees, we do not yet know if they develop in synchrony. This is an important point in the context of the theory outlined in this chapter because if chimpanzees develop representational capacities for some domains (such as objects) but not others (such as mental states), then it may have important implications for the scope of their self-concept. What if, for example, chimpanzees are only able to represent their physical or proprioceptive states? That is, what if for them S_i is composed of only $s_{i,\ \mathrm{physical}}$ and $s_{i,\ \mathrm{psychological\ (agency)}}$? Such a representation would still allow them to represent the aspect of their psychology (their agency) that triggers the equivalence relation through the detection of perfect contingency.

In this case, we would still be warranted in claiming that their capacity to recognize themselves in mirrors reflects (is allowed by) the presence of a self-concept, as Gallup (1970) originally speculated. However, in this case his later theory about the scope of that self-concept would be incorrect (Gallup, 1982).

Finally, what about chimpanzees' conceptions of themselves as entities with a past and a future? Gallup (1982) speculated that the presence of a self-concept in chimpanzees (as indicated by their ability to recognize themselves in mirrors) left open the possibility that they might be able to "begin formulating questions about themselves in relation to historical as well as future events" (p. 242). From the view offered in this paper, the chimpanzee's capacity for such autobiographical memory and temporal projection depends directly upon its capacity to construct higher-order representations. As of yet, we have no definitive evidence concerning its abilities in this arena (see reviews by Cheney & Seyfarth, 1990; Whiten, 1993; Povinelli, 1993, in press; Tomasello & Call, in press). But this does not mean that we should conclude, as has Fraser (1987), that humans are the only species able to conceive of time far removed from the present. The extent of overlap in homologous cognitive developmental pathways among humans and their nearest relatives remains an open, empirical question. Thus, conducting explicit tests of self-recognition with chimpanzees using delayed feedback is a high priority for our own research program. Ultimately, such research will allow us to take a first step toward discovering if chimpanzees, like us, appreciate that they are unique, unduplicated selves trapped in an irreversible arrow of time.

NOTES

1. "Add to this character [of the present and distant selves belonging together] the farther [sic] one that the distant selves appear to our thought as having for hours of time been *continuous* with each other, and the most recent ones of them continuous with the Self of the present moment, melting into it by slow degrees; and we get a still stronger bond of union. And we think we see an identical bodily thing when, in spite of changes in structure, it exists continuously before our eyes, or when, however interrupted its presence, its quality returns unchanged; so here we think we experience an identical *Self* when it appears to us in an analogous way. Continuity makes us unite what discontinuity might otherwise separate; similarity makes us unite what discontinuity might hold apart" (James, 1890/1950, p. 334; italics in original).

2. In a recent Master's thesis, Thomas Suddendorf (1994) makes a similar argument concerning the relation of metarepresentation and temporal projection into the past and future. Like this model, he sees this capacity as opening up the possibility for mental time travel. He also speculates on the question of whether chimpanzees develop this capacity. In contrast, I remain skeptical about the scope of the chimpanzee's theory of mind (Povinelli, 1993, in press). Although our ideas are quite similar in places, they

have been derived independently. Rather than revise this manuscript to reflect his exposition, I refer the reader to his work directly.

3. "The various members of the collection [of the present and distant selves] are felt to belong with each other whenever they are thought at all. The animal warmth, etc., is their herd-mark, the brand from which they can never more escape. It runs through them all like a chaplet and makes them into a whole, which we treat as a unit, no matter how much in other ways the parts may differ *inter se*" (James, 1890/1950, p. 334).

4. In a later attempt to distinguish preconceptual forms of self-knowledge from the conceptual kind of knowledge needed for self-recognition in mirrors, Gallup (1977b) argued for a distinction between "self-sensation" and "self-perception." However, following Butterworth (1992), Gallup (1991) recently adopted a terminology more consistent with that used by developmental psychologists: "self-perception" versus "self-conception." Despite this, his distinction has always been between knowledge about the self coded in terms of proprioceptive and kinesthetic feedback versus a concept of self.

5. There have been numerous demonstrations of this sensitivity to contingency, ranging from demonstration that infants are sensitive to reciprocal behavioral patterns on the part of caregivers, to demonstration that infants detect the contingency between their own (visually obscured) movements and live feedback of that image on a video monitor (e.g., Bahrick & Watson, 1985). Meltzoff (1990) has shown that 14-month-olds are sensitive to others who perform actions contingent with their own, and are especially sensitive to those who imitate the exact form of their behavior. From this kind of evidence it is possible to conclude that by very early in ontogeny, the infant is able to detect actions that are contingent with its own. The reason for this sensitivity is unknown, although detection of contingency in general plays a fundamental role in theories of the simplest forms of animal and human learning (Rescorla, 1967).

6. To my knowledge, two studies have exposed children to playbacks of their previously recorded visual images for the purpose of assessing self-recognition. Brooks-Gunn & Lewis (1984) recorded and coded the responses (affect, interest, imitation) of infants to prerecorded images of themselves or a same-sex age-mate. Zazzo (1982) also presented playbacks of images to young preschoolers. Apparently neither of these studies explicitly assessed self-recognition using a mark test.

7. Nelson (1991) has used Weist's (1986) four-stage model of the temporal systems that young children display from 18 months to 4 years to account for the linguistic referents to temporal events collected from the crib monologues of the child Emily. This model establishes increasingly sophisticated relations between speech time (ST, the here and now), event time (ET), and reference time (RT). The system begins at a point when ET and RT are fixed at ST. This implies that very young children of this age are able to talk about events not localized at the present. The final stage, a free RT system, is not achieved until about 3.5 to 4 years and is characterized by an ability to distinguish RT from ST, and ET from both RT and ST (Weist, 1986). Although it has primarily been used as a means of analyzing linguistic utterances, an analysis of the kind of temporal-conceptual distinctions that become possible in this final free-RT system reveals important underlying parallels between Weist's (1986) four-stage model, Nelson's (1991) elaboration of it, and the model I have developed here.

ACKNOWLEDGMENTS

The theoretical ideas for this chapter were developed during my participation in meetings of the Infant Intentionality Group at Yale University during the spring of 1991. Gordon Gallup, Helen Perilloux, Timothy J. Eddy, Mark Povinelli, and Anthony Maida offered helpful discussions and advice on versions of the manuscript. This work was supported by National Institutes of Health Grant No. RR-03583-05 to the New Iberia Research Center and National Science Foundation Young Investigator Award SBR-85458111 to D.J.P. Address correspondence to Daniel J. Povinelli, Laboratory of Comparative Behavioral Biology, USL-New Iberia Research Center, 4401 W. Admiral Doyle Dr., New Iberia, LA 70560. Telephone: (318) 365 2411, FAX: (318) 373 0057.

REFERENCES

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

Anderson, J.R. (1984a). Monkeys with mirrors: Some questions for primate psychology. *International Journal of Primatology, 5,* 81–98.

Anderson, J.R. (1984b). The development of self-recognition: A review. *Developmental Psychobiology, 17,* 35–49.

Anderson, J.R. (1986) Mirror-mediated finding of hidden food by monkeys (*Macaca tonkeana* and *M. fascicularis*). *Journal of Comparative Psychology, 100,* 237–242.

Anderson, J.R., & Roeder, J.J. (1989). Responses of capuchins monkeys (*Cebus apella*) to different conditions of mirror-image stimulation. *Primates, 30,* 581–587.

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

Bahrick, L.E., & Watson, J.S. (1985). Detection of intermodal proprioceptive-visual contingency as a potential basis for self-perception in infancy. *Developmental Psychology, 21,* 963–973.

Baillargeon, R. (1987). Young infants' reasoning about the physical and spatial properties of a hidden object. *Cognitive Development, 2,* 655–664.

Baldwin, D.A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology, 29,* 832–843.

Bartsch, K. & Wellman, H. (1995). *Children talk about the mind.* Oxford, Oxford University Press.

Benhar, E.E., Carlton, P.L., & Samuel, D. (1975). A search for mirror-image reinforcement and self-recognition in the baboon. In S. Kondo, M. Kawai ,& S. Ehara (Eds.), *Contemporary primatology: Proceedings of the 5th international congress of primatology,* (pp. 202–208). New York: Karger.

Bertenthal, B.I. & Fischer, K.W. (1978). Development of self-recognition in the infant. *Developmental Psychology, 14,* 44–50.

Bigelow, A.E. (1981). The correspondence between self- and image-movement as a cue to self-recognition for young children. *Journal of Genetic Psychology, 139,* 11–26.

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]. *Schweizerische Zeitschrift für Psychologie, 47,* 147–159.

Brownell, C.A., & Carriger, M.S. (1990). Changes in cooperation and self-other distinction during the second year. *Child Development, 61,* 1164–1174.

Butterworth, G. (1992). Origins of self-perception in infancy. *Psychological Inquiry, 3,* 98–109.

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology, 9,* 55–72.

Calhoun, S., & Thompson, R.L. (1988). Long-term retention of self-recognition by chimpanzees. *American Journal of Primatology, 15,* 361–365.

Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world.* Chicago: University of Chicago Press.

Eddy, T.J., Gallup, G.G., Jr. & Povinelli, D.J. (1995). Development of the ability of chimpanzees (*Pan troglodytes*) to distinguish mirror-images of self from video images of others. Manuscript submitted for publication.

Fadil, C.A., Moss, L.E., & Bahrick, L.E. (1993, March). *Infants' visual recognition of their own faces.* Poster presented at the 1993 Meeting of the Society for research in Child Development, New Orleans, LA.

Flavell, J.H. (1988). From cognitive connections to mental representations. In J.W. Astington, P.L. Harris, & D.R. Olson, (Eds.), *Developing theories of mind* (pp. 244–267). Cambridge: Cambridge University Press.

Forguson, L., & Gopnik, A. (1988). The ontogeny of common sense. In J.W. Astington, P.L. Harris, and D.R. Olson (Eds.), *Developing theories of mind.* (pp. 226–243). Cambridge: Cambridge University Press.

Fraser, J.T. (1987). *Time the familiar stranger.* Amherst, MA: University of Massachusetts Press.

Gallup, G.G., Jr. (1970). Chimpanzees: Self-recognition. *Science, 167,* 86–87.

Gallup, G.G., Jr. (1975). Toward an operational definition of self-awareness. In R.H. Tuttle (Ed.), *Socio-ecology and psychology of primates* (pp. 309-341). The Hague, The Netherlands: Mouton.

Gallup, G.G., Jr. (1977a). Absence of self-recognition in a monkey (*Macaca fascicularis*) following prolonged exposure to a mirror. *Developmental Psychobiology, 10,* 281–284.

Gallup, G.G., Jr. (1977b). Self-recognition in primates: A comparative approach to the bidirectional properties of consciousness. *American Psychologist, 32,* 329–338.

Gallup, G.G., Jr. (1982). Self-awareness and the emergence of mind in primates. *American Journal of Primatology, 2,* 237–248.

Gallup, G.G., Jr. (1994). Self-recognition: Research strategies and experimental design. In S. Parker, R. Mitchell, & M. Boccia (Eds.), *Self-awareness in animals and humans.* (pp. 35–50). Cambridge: Cambridge University Press.

Gallup, G.G., Jr., McClure, M.K., Hill, S.D., & Bundy, R.A. (1971). Capacity for self-recognition in differentially reared chimpanzees. *The Psychological Record, 21,* 69–74.

Gallup, G.G., Jr., Povinelli, D.J., Suarez, S.D., Anderson, J.R., Lethmate, J., & Menzel, E.W. (in press). Further reflections on self-recognition in primates. *Animal Behaviour.*

Gallup, G.G., Jr., Wallnau, L.B., & Suarez, S.D. (1980). Failure to find self-recognition in mother-infant and infant-infant rhesus monkey pairs. *Folia Primatologica, 33,* 210–219.

Gopnik, A. (1982). Words and plans: Early language and the development of intelligent action. *Journal of Child Language, 9,* 303–318.

Gopnik, A., & Slaughter, V. (1991). Young children's understanding of changes in their mental states. *Child Development, 62,* 98–110.

Guillaume, P. (1971). *Imitation in children*, 2nd Ed. Chicago: University of Chicago Press. (original work published 1926)

Itakura, S. (1987a). Use of a mirror to direct their responses in Japanese monkeys (*Macaca fuscata fuscata*). *Primates, 28,* 343–352.

Itakura, S. (1987b). Mirror guided behavior in Japanese monkeys (*Macaca fuscata fuscata*). *Primates, 28,* 149–161.

James, W. (1950). *The principles of psychology*, New York: Dover. (original work published 1890)

Johnson, C.B. (1983). Self-recognition in infants. *Infant Behavior and Development, 6,* 211–222.

Kagan, J. (1981). *The second year: The emergence of self-awareness.* Cambridge, MA.: Harvard University Press.

Langer, J. (1986). *The origins of logic: One to two years.* New York: Academic Press.

Leslie, A. (1987). Pretense and representation: Origins of "theory of mind." *Psychological Review, 94,* 412–426.

Lethmate, J., & Ducker, G. (1973). Untersuchungen am ebsterkennen im spiegel bei orangutans einigen anderen affenarten. [Self-recognition by orangutans and some other primates.] *Zeitschrift fur Tierpsychologie, 33,* 248–269.

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

Lin, A.C., Bard, K.A., & Anderson, J.R. (1992). Development of self-recognition in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 106,* 120–127.

Lillard, A.S. (1993). Pretend play skills and the child's theory of mind. *Child Development, 64,* 348–371.

Loveland, K.A. (1986). Discovering the afffordances of a reflecting surface. *Developmental Review, 6,* 1–24.

Macnamara, J. (1982). *Names for things.* Cambridge, MA: MIT Press.

Maestripieri, D., Schino, G., Aureli, F., & Troisi, A. (1992). A modest proposal: Displacement activities as an indicator of emotions in primates. *Animal Behaviour, 44,* 967–979.

Marchal, P., & Anderson, J.R. (1993). Mirror-image responses in capuchin monkeys (*Cebus apella*): Social responses and use of reflected environmental information. *Folia Primatologica, 61,* 165–173.

Meltzoff, A.N. (1990). Foundations for developing a concept of self: The role of imitation in relating self to other and the value of social mirroring, social modeling, and self-practice in infancy. In D. Cicchetti & M. Beeghly (Eds.), *The self in transition: Infancy to childhood,* (pp. 139–164). Chicago: University of Chicago Press.

Meltzoff, A.N., & Moore, M.K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198,* 75–78.

Mitchell, R.W. (1993). Mental models of mirror-self-recognition: Two theories. *New Ideas in Psychology, 11,* 295–325.

Nelson, K. (1989). Monologue as the linguistic construction of the self in time. In K. Nelson (Ed.), *Narratives from the crib* (pp. 284–308). Cambridge, MA: Harvard University Press.

Nelson, K. (1991). The matter of time: Interdependencies between language and thought in development. In S.A. Gelman & J.P. Byrnes (Eds.), *Perspectives on language and thought* (pp. 278–318).

Nelson, K. (1993). The psychological and social origins of autobiographical memory. *Psychological Science, 4,* 1–8.

Olson, D.R. (1993). The development of representation: The origins of mental life. *Canadian Psychology, 34,* 1–14.

Olson, D., & Campbell, R. (1993). Constructing representations. In C. Pratt & A.F. Garton (Eds.), *Systems of representation in children: Development and use* (pp. 11–26). New York: John Wiley & Sons.

Parker, S.T. (1991). A developmental approach to the origins of self-recognition in great apes. *Human Evolution, 6,* 435–449.

Perner, J. (1991). *Understanding the representational mind.* Cambridge, MA: MIT Press.

Povinelli, D.J. (1989). Failure to find self-recognition in Asian elephants (*Elephas maximus*) in contrast to their use of mirror cues to discover hidden food. *Journal of Comparative Psychology, 103,* 122–131.

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

Povinelli, D.J. (in press). 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., & Eddy, T.J. (in press). What young chimpanzees know about seeing. *Monographs of the Society for Research in Child Development.*

Povinelli, D.J., Landau, K., & Perilloux, H.K. (in press). Self-recognition in young children using delayed versus live feedback: Evidence of a developmental asynchrony. *Child Development.*

Povinelli, D.J., Rulf, A.R., Landau, K., & Bierschwale, D.T. (1993). Self-recognition in chimpanzees (*Pan troglodytes*): Distribution, ontogeny, and patterns of emergence. *Journal of Comparative Psychology, 107,* 347–372.

Priel, B., & de Schonen, S. (1986). Self-recognition: A study of a population without mirrors. *Journal of Experimental Child Psychology, 41,* 237–250.

Rescorla, R.A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review, 71,* 71–80.

Robinson, J.A., Connell, S., McKenzie, B.E., & Day, R.H. (1990). Do children use their own images to locate objects reflected in a mirror? *Child Development, 61,* 1558–1568.

Schulman, A.H., & Kaplowitz, C. (1977). Mirror image response during the first two years of life. *Developmental Psychobiology, 10,* 133–142.

Suarez, S.D., & Gallup, G.G., Jr. (1981). Self-recognition in chimpanzees and orangutans, but not gorillas. *Journal of Human Evolution, 10,* 175–188.

Suddendorf, T. (1994). *Discovery of the fourth dimension: Mental time travel and human evolution.* Unpublished Master's thesis. University of Waikato, Hamilton, New Zealand.

Swartz, K.B., & Evans, S. (1991). Not all chimpanzees show self-recognition. *Primates, 32,* 483–496.

Tomasello, M. & Call, J. (in press). Social cognition of monkeys and apes. *Yearbook of Physical Anthropology, 37.*

Weist, R.M. (1986). Tense and aspect. In P. Fletcher and M. Garman (Eds.), *Language acquisition,* 2nd Edition (pp. 356–374). Cambridge: Cambridge University Press.

Whiten, A. (1993). Evolving theories of mind: The nature of nonverbal mentalism in other primates. In S. Baron-Cohen, H. Tager-Flusberg, D. Cohen, & F. Volkmar (Eds.), *Understanding other minds,* (pp. 367–396). Oxford: Oxford University Press.

Zahn-Waxler, C., & Radke-Yarrow, M. (1982). The development of altruism: Alternative research strategies. In N. Eisenberg (Ed.), *The development of prosocial behavior* (pp. 109–137). New York: Academic Press.

Zazzo, R. (1982). The person: Objective approaches. In W.W. Hartup (Ed.), *Review of child development research: Vol. 6*(pp. 247–290). Chicago: University of Chicago Press.