8

Extending Self-Consciousness Into the Future

John Barresi
Dalhousie University

As adults we have little difficulty thinking of ourselves as mental beings extended in time. Even though our conscious thoughts and experiences are constantly changing, we think of ourselves as the same self throughout these variations in mental content. Indeed, it is so natural for adults to think this way that it was not until the 18th century—at least in Western thought—that the issue of how we come to acquire such a concept of an identical but constantly changing self was first recognized as a problem that required an explanation. Philosophical discussion of this issue was initiated when John Locke (1694/1975) proposed a notion of personal identity and selfhood based on consciousness:

> For since consciousness always accompanies thinking, and 'tis that, that makes every one to be, what he calls *self*; and thereby distinguishes himself from all other thinking things, in this alone consists *personal Identity*, i.e. the sameness of a rational Being: And as far as this consciousness can be extended backwards to any past Action or Thought, so far reaches the Identity of that *Person*; it is the same *self* now it was then; and 'tis by the same *self* with this present one that now reflects on it, that that Action was done. (p. 335)

According to this view, we are the same self insofar as we can consciously accept as our own not only those mental and physical acts that we perform *now* but also those acts done in the *past*, that we can

recollect as our own. This Lockean view of self based on consciousness was an advance beyond Descartes's (1641/1984) notion that the self was an immaterial soul or mind, whose identity was guaranteed not by being conscious of itself but by being a substance distinct from the body. However, Locke's (1694/1975) more empirical approach to personal identity and self raised a new question: How do we come to form the notion of this conscious self that is extended in time? Some progressive followers of Locke—such as Hume (1739/1888)—believed that our notion of an *identical* extended self is, in fact, a fiction. On this view there is no *actual* mental self that is identical through time, merely an *illusion* of such an identical self. Although this suggestion was an important advance toward a strictly psychological account of the origins of our concept of self, Hume did not focus on the developmental problem of how or when this fictional notion of an identical but extended mental self is formed. It was not until 1805, when William Hazlitt, the well-known English Romantic critic, published *An Essay on the Principles of Human Action* (Hazlitt, 1805/1969; Martin & Barresi, 1995, 2000), that a truly developmental account of the formation of the concept of an extended self was first proposed.

Unlike Locke, who focused his account of the extended self on our self-conscious recollection of the past, Hazlitt was more interested in our relation to the future. Why should we be concerned about a self, whom we might become in the future? In what sense is this future self, with its different motives and goals, the *same* self that we are now and have been in the past? And how do we come to believe that it is the same self as the present and past self? In other words, how do we form a notion of a changing mental self that extends not only into the past but also into the future?

In attempting to answer these questions, Hazlitt (1805/1969) put forward a fascinating hypothesis concerning the development of the concept of a self that extended into the future. Hazlitt felt that, unlike our relationship to our present and past selves, which depend on causal relations of sensation and memory, our relationship to our future selves can depend only on *imagination*, because we can have no causal relations working backward from our future to our present selves. For Hazlitt, an important consequence that follows from this assumption is that we can have no *necessary* interest in *any* future self. All interest in our future selves must be mediated by imagination, which is our only connection with the future. However, imagination, which connects us to the future interests or motivations of ourselves, can operate no differently, *at least initially*, when imagining our *own* future interests than it can when imagining the future interests of *another* person. Hence, he concluded that, insofar as we make voluntary choices and reflect on future interests or desires, we are not "naturally" self-interested. All self-interestedness involving our future selves must be acquired through experience. Before

such self-interest is acquired we have no more present concern about the interests or desires of our future selves than about those of any other person.

In addition to his hypothesis about the origins of a future-oriented *extended self* Hazlitt also recognized, as a developmental problem, the issue of how we come to know ourselves as selves *at all*. He believed that one could know one's self as a self only if one could discriminate one's own thoughts or mental states from those of another and that making this discrimination required knowing that another individual was also a *self*. Learning to distinguish one's own emotions, motives, and other mental states is an achievement that implies forming a distinction between self and other, where previously no such distinction was made.

Hazlitt's ontogenetic perspective on the origins of self-concepts has found a parallel in recent developmental studies. His suggestion about the origin of the concept of the self receives support from some recent theories that claim that infants become aware of shared mental states before distinguishing their own mental states from those of others (e.g., Barresi, 2000; Barresi & Moore, 1993, 1996). For instance, in the area of emotions it is only after the first year of life that infants begin to distinguish their own original emotions from those that they experience contagiously from others (Hoffman, 1976). Furthermore, evidence suggests that it is not until the end of the second year of life that the child begins to exhibit clear signs of self-consciousness (Barresi & Moore, 1996, for a review). This evidence includes self-recognition in a mirror task; self-conscious emotions, such as embarrassment, that imply a reflexive self-awareness; and linguistic differentiation of self and other. However, this acquired self-consciousness involves only the here and now. This suggests that the infant has achieved only a concept of a *present* or *now self*, which—because of the infant's lack of awareness of its temporal and changing properties—is treated as if it were an unchanging, or permanent, self.

It is not until children are past 3 and closer to 4 years old that they begin to show evidence of a self-consciousness that extends through time. Most of the research that provides evidence of this shift in self-consciousness has focused on the child's ability to remember past events in a serial fashion and to integrate the past with the present (cf. Gopnik & Slaughter, 1991; Nelson, 1992, chap. 2, this volume; Perner, 1992, chap. 10, this volume; Perner & Ruffman, 1995; Povinelli, 1995, chap. 5, this volume). For instance, the child at this time acquires the ability to recognize him- or herself in videotapes taken earlier in a way that indicates an understanding of how past events involving the self relate to one's self in the present. This videotape task is particularly interesting because of its analogy to the mirror recognition task. Povinelli has shown that 3-year-olds are perfectly competent at recognizing themselves in the replay of a recent video (Povinelli, 1995; chap. 5, this volume). However,

young 3-year-olds fail to appreciate the relation between stickers placed on their heads in a recent video and the possibility that the stickers might still be there. If they remove the stickers on seeing a video of the stickers being placed on their heads, it does not seem to matter to them whether the video was taken a few minutes earlier or a week earlier. Older 3-year-olds are more sensitive to this variation in video timing and more likely overall to remove the sticker immediately on seeing the video. Although this task alone might not indicate the development of a temporally extended self-consciousness, it is during this same period that the child begins to show autobiographical knowledge of past events and becomes capable of recalling past mental states of self. Taken together, these results suggest that a shift in self-consciousness has occurred, from one involving a permanent "present" or "now" self to one involving a concept of a temporally extended, ever-changing self—what Neisser (1988) calls the *extended self*.

To date, the bulk of this research on the temporally extended self has focused on extensions of self into the past, much as Locke originally posed the problem of personal identity. It has not dealt with extensions of self into the future. In this chapter I describe some research that focuses on our relations to future selves. In this research we pursued Hazlitt's suggestion that the use of imagination with respect to future mental states of self and other is—at least initially—symmetrical, and that, as a result, future-oriented prudence and future-oriented altruism ought to develop concurrently. Furthermore, we took the position that Hazlitt's theory of imagination is analogous to current simulation accounts of theory of mind (e.g., Harris, 1991), which implies that the development of future-oriented prudence and altruism ought to emerge around 4 years of age. After briefly describing these experiments, which have been more fully reported elsewhere (Moore, Barresi, & Thompson, 1998; Thompson, Barresi, & Moore, 1997), I use some findings from these experiments as a basis from which to develop a theoretical account of the emergence of extended self-consciousness.

EXTENDING THE SELF INTO THE FUTURE

The basic paradigm that we used in this research was Mischel's delay-of-gratification choice procedure (Mischel, 1974; Thompson et al., 1997). However, the rewards were stickers that could be placed in a sticker book immediately or later, after the sticker choice activity was finished. Furthermore, the children who were tested were asked to make decisions that would affect another person as well as themselves.

On the basis of Hazlitt's view that imaginative projection into a future self and other are analogous, and that children have a natural sympathy toward others as well as toward self, we built the study design

to measure how children would acquire concepts about the future, when both self and other were involved. The general hypothesis that we entertained in the first experiment was that younger children would prefer immediate rewards to delayed rewards and that this would occur for choices involving the self alone as well as involving self and other.

The four choice conditions about which the child had to make decisions were:

1. *Shared-gratification condition without cost*: The child had to choose between having one sticker for the self now, or one sticker for the other person as well as the self now.

2. *Shared-gratification condition with cost*: The child had to choose between having two stickers for the self now or one sticker for the other and one for the self now.

3. *Delay-of-self-gratification condition (or future-oriented prudence)*: The child had to choose between one sticker now for the self or two stickers later for the self.

4. *Delay-of-shared-gratification condition (or future-oriented altruism)*: The child had to choose between one sticker now for the self or one sticker each later for the self and the other.

Each child received three blocks of choices; each block had one of each of these four choice conditions.

We anticipated that all children (3-, 4-, and 5-year-olds) would prefer having both self and other receive rewards over the self alone because it has been shown that even 2-year-olds exhibit empathy toward others. However, this preference was expected to extend only to present sharing and not to future sharing for young children. We anticipated that only the older children would be willing to forgo a present reward for the self in favor of a larger reward in the future and that this preference would apply to conditions involving the self alone, thus implying prudence, as well as self and other, thus implying altruism.

As shown in Fig. 8.1, these hypotheses were confirmed by the results of the experiment. Out of three trials of the same type, all groups showed a preference for shared reward in the present over self-alone rewards in the present, and this did not vary as a function of age. More important, 3-year-olds, but not 4- or 5-year-olds, preferred present to future rewards, whether this involved increased future rewards for self alone or for self and other. Furthermore, there was a correlation of responses between these two delayed reward choice conditions for 3-year-olds, but not for the older children.

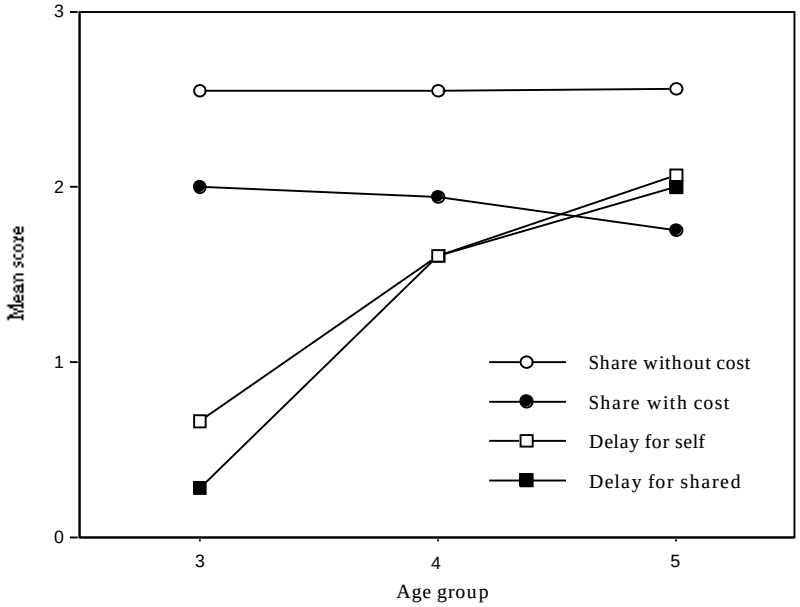

FIG. 8.1 Mean share or delay scores for different choices in sticker-choice task for three age groups in Experiment 1. The choice conditions are: share without cost (one sticker for self now vs. one sticker for each now), share with cost (two stickers for self now vs. one sticker for each now), delay for self (one sticker for self now vs. two stickers for self later), delay for shared (one sticker for self now vs. one each later).

These results support the hypothesis that it was the factor of delay that caused a problem for younger but not for older children. They suggest that 3-year-olds, although capable of dealing with the mental states of self and other in the present, are not able to make decisions that would be preferred by their own future self or the future self of another person. In terms of consciousness, the 3-year-olds show evidence of being conscious of their own immediate or present desires as well as of the supposed immediate desires of the other, but they fail to have any empathy for their own future self or the future self of the other. Apparently, they are not yet conscious of the self or other as extended in time, with desires that can conflict between selves from different times.

I want now to discuss two other experiments that my colleagues and I conducted (Moore et al., 1998) that bear on the issue of how 3-year-olds come to form a concept of a self that extends through time. It seemed to us that two mechanisms might be involved: First, children's ability to represent to themselves mental states of self and other that conflict with their own current mental states, and second, children's ability to engage in executive decision making that requires inhibition of responses based on the children's own current desires in order to make future-oriented choices. Both of these abilities have been shown to develop during the fourth year of life.

In both of the follow-up experiments we included the sticker-choice task, and replicated the two conditions from the first experiment that involved the future, but we varied the other two conditions. In the other two conditions of the second experiment the children were asked to make choices for the other person rather than for themselves, but this involved the same two future choice conditions—prudence and altruism—only now for the other person (see Choice Conditions column in Table 8.1). In addition, in this experiment we assessed the child's ability to perform on three theory-of-mind tasks (False Belief; Representational Change, and Desire Change). In all three tasks the children (old 3-year-olds and young 4-year-olds) were first exposed to a situation in which they initially believed or desired one thing about an object and, later, believed or desired a conflicting thing about the same object. In the self-change tasks (Representational Change and Desire Change) they were asked about their former belief or desire. In the other belief task (False Belief) they were asked to guess what another person naïve to the original situation would think. The critical aspect of all three tasks is that the children's own current mental states after they acquired knowledge of the object, or fulfilled their initial desire, would conflict with their prior mental states, which—in the case of belief—would be the normal naïve state of another person.

The results of this experiment replicated the findings showing a developmental change in children's ability, during the fourth year, to

TABLE 8.1
Experiment 2 Choice Conditions and Correlations With Composite Measures of Theory of Mind

	Choice Condition	***Theory-of-Mind Tasks***
Self	One for self now vs. 2 for self later	
	One for self now vs. 1 for each later	Correlated with choice condition in 4-year-olds
Other	One for other now vs. 2 for other later	Correlated with choice condition in 4-year-olds
	One for other now vs. 1 for each later	Correlated with choice condition in 4-year-olds

deal with the future in the choice conditions involving the self. Again, we found a correlation between prudence and altruism choices involving the future in the group of old 3-year-olds tested in this experiment. We also found that all the children had trouble dealing with the choice task when it involved choices for another person rather than for the self. It is interesting, however, that we found that the young 4-year-old's performance on this task was correlated with their ability at theory-of mind-tasks (see Table 8.1, which shows a similar correlation for the future-oriented altruism condition).

What these results suggest is that the ability to do theory-of-mind tasks may be a necessary skill for children to be able fully to imagine another person's choice situation, when that involves the future. As I argue later, however, these results can also be interpreted as indicating that it is only 4-year-olds who have a sufficiently rich and stable concept of the extended self to be able to make choices requiring simulation of the point of view of the extended self of another person. By contrast, the ability to make such choices for the self, and the origins of the extended self, seems to develop at an earlier time and may even precede the development of a theory of mind. But how does the child come to form this concept of his or her own self as extended? And how might it contribute to the development of a theory of mind? I believe our third experiment (Moore et al., 1998) speaks to this issue.

In the third experiment we (Moore et al., 1998) again replicated the two future-choice conditions previously tested but added two new future choice conditions: a choice condition involving a decision between one for each now or two for the self later and another requiring a decision between two for the other now or one for the self later (see Table 8.2). The children (young 3-year-olds, old-3-year-olds, and young-4-year-olds) were also tested on the "windows" task, an executive function task that normal children learn to pass around the middle of the fourth year (Russell, Mauthner, Sharpe, & Tidswell, 1991). The windows

TABLE 8.2
Experiment 3 Choice Conditions and
Correlations With Executive Function Task

	Choice Condition	***Executive Function Task***
Old	One for self now vs. 2 for self later (conflict)	Correlated with choice condition in young 3-year-olds
	One for self now vs. 1 for each later (conflict)	Correlated with choice condition in young 3-year-olds
New	One for each now vs. 2 for self later (conflict)	Correlated with choice condition in young 3-year-olds
	Two for other now vs. 1 for self later (no conflict)	

task requires the child to learn a rule according to which they have to point to an empty box in order to acquire the reward in another similar box. Russell et al. (1991) showed that all children in this age range can learn the rule and follow it when the contents of the boxes are not visible at the time of pointing. When they are switched to boxes with windows, however, the younger children fail to generalize their performance. It seems that they cannot inhibit their natural tendency to point to the box with the visible reward, although they are able, verbally, to describe the rule requiring them to point at the empty box in order to obtain the reward.

We replicated these findings in the present experiment. Most young 3-year-olds, but few older children, failed at the windows task. More important, we found that only for the young 3-year-olds did performance on the windows task correlate with their performance on any of the sticker-choice tasks (see Table 8.2). For the three conditions in which the child had to make a choice involving conflict between current and future rewards for self, their preference for the delayed but larger future rewards correlated with their ability to point to the empty window in order to obtain rewards in the windows task. The one choice condition that didn't correlate with performance on the windows task was the condition in which the choice was between a current reward for the other and a future reward for self—a condition that did not produce conflict and in which most children chose the future reward for self.

EXTENDING SELF-CONSCIOUSNESS

The results of these three experiments provide evidence that children acquire the ability to deal with future mental states of self and other in choice situations involving prudence and altruism during the age-4 transition. These results also suggest the possibility that there are at least two independent skills involved in acquiring this ability to deal with future mental states of self and other. The first of these skills is an executive capacity that allows the child to inhibit responding to current desires in order to obtain future rewards. This skill, which appears to develop during the early part of the fourth year, directly affects all decisions involving conflict between current and future-oriented desires of self. The second skill is the ability to imagine or think about the future desires of another self or person as well as the ability to think about another person's decision processes in choice situations similar to those that the younger child can already handle for him or herself. This latter skill arises somewhat later—possibly as late as in young 4-year-olds. I would now like to speculate on the meaning of these results, in conjunction with other findings relating to the age-4 transition, for the development of an extended conception of self.

What is remarkable about the changes that occur around the fourth year of life with respect to the self is that children become able to treat their own past and future selves as if these selves were other persons, with points of view different from their current self. With respect to the past, children can now remember having had experiences that differ from those that they have at present. With respect to the future, as the present research suggests, they can anticipate having experiences that they do not have at present. As Hazlitt's (1805/1969) hypothesis suggests, this is not accomplished by imagining being the very same self that one is now but located in the past or future. Instead, it is accomplished by imagining being an altogether different self, who exists at a different time, with a different perspective on things. Hazlitt suggested that there is an asymmetry between past and future with respect to the present self, where the past involves memory and a causal relation, whereas the future requires imagination. But Hazlitt, who often mentioned the use of imagination with respect to the past as well as the future, was not always consistent in his view on how one relates to "other" selves in the past as well as in the future. Hence, a more detailed analysis than he presented is required.

Consider first the future. A person might anticipate having certain experiences in the future without having to consider possible differences between his or her current and future selves. In such a case the role of imagination is limited to simulating a future situation without changing one's attitude about that situation. For instance, in the present-choice task children might imagine what it would be like to make a choice that results in immediately receiving a sticker to place in their sticker book. In such a case the children's desire for the sticker will be immediately gratified, thus their imagining that gratification following the choice requires no adjustment in their desires—no imagining of different desires of future selves. On the other hand, they may try to imagine themselves in a future situation, where their own viewpoint is itself changed in that context from the viewpoint that they currently have. Thus, children in the choice task might try to imagine getting many stickers at the end of the sticker game if they make the delayed choices now. In this case it won't do to imagine making a choice that is immediately gratified with a sticker or several stickers to put in the book. They must imagine making a choice that leads immediately to no stickers and hence, not satisfy their immediate desire, but which nevertheless will eventually make them happy—but only after they, themselves, have changed. It is only in this latter kind of situation that one is truly imagining a future self as an "other." It is only in this situation that a person makes a distinction between his or her own current point of view, his or her *now* self, who will not be immediately gratified by the decision, and a *future* self, who will be so gratified yet with whom he or she might still be able to identify.

The same distinction applies to one's possible relations to the past. On the one hand, one can remember particular events and happenings as if one were there but as the same self that one is at the moment. For instance, children in the sticker task might remember getting stickers on previous occasions, after making a nondelayed choice, and of not receiving a sticker when they make a delayed choice. Now, if they are still in the midst of making choices, their attitude to these two types of situation won't have changed very much. They will remember the nondelayed choices followed by gratification with a more positive feeling than the delayed choices. On the other hand, one can recollect having had certain experiences, from the point of view of a past self, while at the same time having a distinct current personal viewpoint that differs from this past self. This could happen to some of the older children at the end of the sticker game, when they receive the delayed rewards. At that time, they might recall originally being unhappy about their delayed choices and happy about their nondelayed choices but now feeling happy about the delayed choices and unhappy about the nondelayed choices. In this situation the person or child distinguishes how things seem to him or her now from how they seemed to him or her in the past. Whereas in the past he or she had one attitude toward the situation, now he or she can recall having had that attitude but not currently feeling the same way about the situation.

To appreciate the processes involved in memory of past events it is important to consider several distinctions. One distinction is that between reliving an event and remembering it. In reliving an event, one re-experiences the past from the point of view of the past self, entirely forgetful of the present self. In the example just considered, this would involve re-experiencing the choice made in the past from the point of view that one had in the past but not noticing that this is a memory, rather than a re-experiencing, of the past event. In remembering it, one recalls what happened in the past, but one realizes that the event occurred in the past while also maintaining one's present point of view. However, remembering also takes several forms. One of these forms happens when one remembers the past event as an occurrence rather than as an experience. If children at the end of the game merely remembered that they previously made very few delayed choices, without remembering their experience of making those choices, they would have this kind of memory of the past. In this case they might now feel sorry about the choices they made then, without remembering that at the time they felt quite differently about the choices. When a person both remembers what happened and how he or she experienced that happening at the time, we say he or she recollects his or her past. This is the kind of self-conscious memory that Locke (1694/1975) described as essential to personal identity. In such a recollection of the past one distinguishes between how one experienced the event then and how it

appears to one now, and one can appreciate both past and present points of view as distinct. Yet both viewpoints are conceived as phases in the personal history of the same temporally extended self. In a sense, recollection involves a kind of empathic re-experiencing of the past while not fully identifying with this past self (see Reed, 1994; also compare Lemmon & Moore, chap. 9, and Perner, chap. 10, this volume). One must distinguish the two selves and not collapse them into either the present self or the past self yet also appreciate that they both belong to the same person, or self, as extended in time. The same distinction holds true with respect to the present and future selves. Perhaps one can imagine being some future self, a sort of "living" this future self. However, for this future self to have an independent reality from one's current self one must distinguish the two, so that they are not confused with each other, yet both are conceived to represent phases of a temporally connected person. In the case of children thinking about future rewards, they must distinguish how they will feel then about their choice now and how they feel now in making that choice. Furthermore, they must identify sufficiently with the future self that they are willing to accept current dissatisfaction in order to gratify the desires of the future self. In effect, they must see the future self not as another conflicting self but as a temporal extension of the current self into the future.

As one sees here there is a perfect symmetry in what is required to distinguish a future and a past self from the self of the present. It should be noted, however, that there is a difference—one that is congruent with Hazlitt's (1805/1969) notion of asymmetry. In the case of the past self, the imagination of that self, although in a sense reconstructive, has prior experience on which to build this reconstruction. In the case of the future self, imagination has to do the work of construction in the absence of specific memorial content. It can use past selves to help imagine a future self, but it cannot use a specific past self to build this imagined future self. Hence, it is, in some sense, surprising that the abilities to deal with past and future selves seem to develop at the same time. As Hazlitt did, one would expect that the child's ability to extend to a past self would precede his or her ability to extend to a future self, but this appears not to be the case (see Lemmon & Moore, chap. 9, this volume). I consider shortly an explanation of why this symmetry may occur.

It might seem at this point that all that is required to form a conception of an extended self is to understand that past and future selves are different from one's present self. More is required, however, and, theoretically, what remains to be dealt with is probably the most crucial part. This is the requirement that one have a conception of the present self as also temporally located—in the now. One must be able to conceive that the present point of view is no different from one in the past or in the future in being a point of view with a temporal locus, with particular experiential properties. Without this differentiation, one's

consciousness of a mental self that transforms through time cannot be conceptualized. As in the case of the past and future selves, to differentiate this now self from the present self one must not "live" in the experience but obtain some psychological distance from it. And, in accomplishing this psychological distance from the now self, one not only backs off from being embedded in the present but also makes possible a conception and consciousness of a self that is extended in time. From the point of view of this extended self-concept the current now self is seen as only a phase of a self that extends from the past through the present and into the future (cf. Nagel, 1970). In a very real sense, the child thus creates a new level of personal unity and self-consciousness, one where the immediate motives and points of view of past, present, and future selves are all treated on a par—each with a right to demand allegiance of the newly created extended self, but none with any special priority over the others.

THE DEVELOPMENT OF THE EXTENDED SELF

How does this conception of the origins of an extended self relate to the present research? One possible interpretation, depicted in Table 8.3, is that there are three phases in the development of an extended self. During the first phase the child must develop sufficient self-consciousness to separate him or herself from his or her current immediate motives and to inhibit responding to a potent stimulus based on immediate desires. What emerges here is an ability to distance oneself from one's immediate desires, by taking up the stance of a reflexively aware self rather than that of an impulse motivated self. During the second phase this reflexively aware self then develops more generally into an extended self and concurrently develops a general theory of

TABLE 8.3
Three Temporal Phases With Approximate Ages
in the Development of an Extended Self.

Young 3s	*Old 3s*	*Young 4s*
Becomes reflexively aware of current mental states in the process of learning to inhibit making responses based on immediate desires; acquires the ability to imagine mental states in conflict with current mental states	Can now use the ability to imagine mental states in conflict with own current states, to form a concept of self as extended, and becomes better at making future choices for self; also begins to form an abstract theory of mind	Becomes efficient at theory of mind and has abstract theory of an extended self, thus can now imagine the future-choice situation for the extended self of another person

mind. Children develop not only a distancing awareness of their own current point of view of reality but also the ability to compare this current point of view with other points of view, including those from their own past and anticipated future as well as those of others. Indeed, even the memory of past mental states may depend on first developing this capacity for reflexive awareness of current mental states as modes of experience or points of view of internal and external objects. These current states of reflective awareness then become the objects of memory, those objects of episodic memory, which tell us how we experienced things in the past (cf. Perner, chap. 10, this volume; Perner & Ruffman, 1995; Tulving, 1985; Wheeler, Stuss, & Tulving, 1997). The current states, from which the child has acquired some psychological distance, might also, through memory and imagination, provide the child with models for how he or she might feel in the future under changing circumstances. Because the development of such a reflexive awareness is a necessary condition for developing a more general notion of ourselves and others as mental beings extended in time, this reflexive awareness would develop somewhat earlier than the more general theory of a mental agent extended in time. The development of such a general theory of an extended mental self would then represent the third phase of development. In partial agreement with these speculations, delay choices were correlated in Experiment 3 with executive function only in young 3-year-olds, whereas delay choices involving others were correlated in Experiment 2 with theory-of-mind tasks only in young 4-year-olds.

However, it is important to note that the two processes—reflexive awareness and the ability to imagine alternative states of mind—that I represent here as being part of the first and second phases of development cannot develop entirely independently of each other. In order to gain some distance from one's current mental state to view it objectively, one must be motivated to reflect on it and have some means by which to begin to conceptualize it. If one attends to the situations in executive function tasks, such as the window task, and the choice task, we can get some idea of what motivates reflection and how "other" minds are sometimes implicated.

In the windows task it is probably an internal conflict within the child, rather than a conflict involving another self, that is at issue (Russell, Jarrold, & Potel, 1994). Even so, it can be understood as a conflict of two "minds" within the child that must become unified. The child has two dispositions to act: one that is habitual, and another that is not. What the child needs to "control" in this situation is his or her habitual response that tends to dominate over the new, nonhabitual, and correct response. To do this, the child must "reflect" on his or her own habit, become aware that there is this inclination to respond in a certain way, and inhibit this response in this particular case. In effect, the child is in two states of mind about how to behave. In one state of mind the child

is reflectively aware of the rule that is to be followed and is prepared to follow it. In the other of state of mind the child is not yet reflectively aware of his or her disposition to act in a certain habitual way and is prepared only to give this habitual response. In order to be able to "follow" the rule, the child must become reflectively aware not only of the rule but also of the habitual response. The child must achieve a unity of mind—that is, a unity of self-consciousness—that can consider both alternatives at once and follow the rule while concurrently inhibiting the habitual response. In doing so the child must become able to reflect on the inclination to respond and to turn it into a choice situation between two possible responses, with the new rule being the choice that is made (cf. Zelazo & Sommerville, chap. 12, this volume).

Now consider the choice task and see how it compares to this process. In the choice task there is again a natural inclination on the part of the child: to request the current reward, regardless of long-term consequences. At least for young children, future rewards can have current value only when they are not in conflict with other current desires. In the third experiment this happened only in the sticker-choice condition in which the children were asked to make a choice between two stickers for the other person now versus one sticker for self. This is a complex choice for children to make, if they fully empathize with the viewpoint of the other person. It is likely, however, that the children take a simpler route in this situation. If they follow self-interest and ignore the interest of the other person they can choose the option for which they get some kind of reward, although a delayed one, over the option for which they get no reward. Probably as a result of such thinking, even young 3-year-olds tend to choose the option for which they will get a delayed reward. The issue of time is not important here for the child because there is no conflict between a present and future self. Both selves prefer getting a reward to getting none: The earlier self can anticipate getting the reward later without having to give up anything important in the present, and the later self can be gratified when he or she receives the reward. Having to wait is just a condition on gratifying the desire to get some kind of reward for the self, so there is no conflict between earlier and later versions of self—the imagined future self and the present self are at one.

However, in the remaining sticker-choice conditions in the third experiment the children must somehow both imagine a future self's point of view, which conflicts with their own current desires, and choose in favor of that future self. Suppose for the moment that the children can actually imagine what the future self's preference would be. Thus, in the future-oriented prudence condition they can imagine getting many stickers at the end of the sticker game if they make the delayed choices now. In doing this they must also imagine making a choice that leads immediately to no stickers, which is in conflict with their own immediate

desire. As a result there is a true conflict between their present and future selves. How do they overcome the conflict and make the choice that will make the future self happy but not gratify the immediate desire? What is required is analogous to what happens in the windows task, but it adds a temporal dimension not found in that task. The children must inhibit their natural inclination to take what they can get now in order to get the reward later. They must give a response in support of the future self over the present self. However, in doing that they must clearly distinguish these two selves and identify with the future self over the present self. The children's task here is to conceptualize the self as a unified self extended in time, distinguishing current feelings from those that they will feel later and not allow their choice to be dominated by current motives alone.

According to current theorizing about the relation among *autonoetic* consciousness (Tulving, 1985), episodic memory, and the extended self, this is just the kind of situation that should promote the development of a conception of an extended self. It is a situation that requires the child to conceive of him- or herself in terms of a temporal dimension, with—in the current context—a present and a future, where the self differs or changes according to time. Thus it is similar to what occurs in episodic memory, which requires a distinction between a past self and a present self, who changed through time. Furthermore, evidence from experiments conducted by Lemmon and Moore (see chap. 9, this volume) shows that the development of the future-oriented extension of the self, and the past-oriented extension of self, are correlated in older 3-year-olds and 4-year-olds. Thus, the child seems to be developing a general abstract representation of his or her self extended in time from the past, through the present, and into the future. And it is here that one may find the reason for this symmetry between acquiring a conception of the self that extends into the past at the same time as a self that extends into the future. What often motivates the development of reflexive self-awareness is some possible reward in the future, but once one is reflexively self-aware of being in a particular state, it leads to a recollective memory of oneself in that state and of the future-oriented choice one made in that state. Thus, in the present-choice task the child at the end of the game can recollect desiring the immediate reward on earlier choice trials. Nevertheless, he or she can also recollect having chosen the delayed reward in order that his or her present—at that time future—self might gain a greater number of stickers, than his or her past—at that time present—self could have received.

Acquiring such a notion of an extended self is no minor accomplishment. It is the means by which children acquire a self-conscious personal identity. What is crucial here is that children begin to think of themselves each as beings with coherently connected mental lives that extend through time. They know how their past choices have

resulted in their present situation, and they know how their present choices will result in their future situations. They no longer live just in the present. They are now persons with a past that they know and a future that they anticipate having, and they view themselves as abstract entities who live through these temporal phases of existence.

THEORY OF MIND AND THE EXTENDED SELF

At about the time that children form a concept of the self as an extended self they also begin to understand other humans as selves or persons extended in time. Indeed, it may be that fully acquiring a theory of mind requires this particular kind of concept of a self or person. In order to understand mental states as representational rather than as presentational of states of the world, or of the organism, it may be necessary to conceive of the individual mind as existing outside of a particular time. If this were the case, as indicated in Table 8.3, it could account for the delay found in the present research between the executive-function and theory-of-mind tasks in their respective correlations to delay choices in the sticker task.

To more clearly understand this possibility consider the actual structure of the theory-of-mind tasks used in the research described earlier and how its structure relates to the extended self. The child in the Belief Change and the Desire Change tasks is required to be able to conceive of him- or herself changing from one state to another state as a function of time and of the events that occur between these two states. A child who lives in the present, with only a sense of an immediate self, but not a sense of an extended self in time, cannot be expected to organize such information about him- or herself. There is an inherent contradiction between the self the child currently experiences him- or herself to be and the past self that the child is requested to describe. This proves to be a difficulty that the child cannot overcome until he or she has a concept of self that extends through time. With such a concept, the child can conceive of him- or herself as a being capable of such contradictions because changes in the self over time allow for them. However, what is at issue here is more than a change in the physical activities of the self, it is the self's nature as a changing mental agent with which the child must become acquainted. Until children can conceive of their mental states as different at different times and as a function of temporally changing situations, they do not actually have a self of which they are conscious, whose very essence is to exist outside of particular times and to change through them. So it is a necessary condition for passing these tasks that the child already has, or be on the way to having an extended self. The same applies to the False Belief task, which requires the child to think about contradictory mental states of another

person. The other individual has to be conceived as having true and false knowledge of the same "fact"—for example, the content of a box of Smarties. What the other person believes to be true now was true at one time (for many other boxes of Smarties) but is now false (for this box). Only a self that can have true and false mental states about the same represented thing, but at different times, can remain the same while also having both of these mental states. So again the child is required to think of a self whose essence is a mental nature that changes as a function of time and situation.

In this research (Moore et al., 1998) the theory-of-mind tasks correlated with the choice task only for young 4-year-olds and, in particular, theory of mind related to the choice task only when the child needed this ability in order to take specific cognizance of another person's mind. Although in this research we did not collect theory-of-mind and executive-function measures in the same experiment, other research shows that theory of mind is correlated with executive function tasks in 3- and 4- year-olds (Carlson, 1997; Hughes, 1998a, 1998b). I think that these results, taken together, are congruent with the idea suggested earlier that theory of mind is an outgrowth of the processes involved in executive decision making (see Table 8.3, Phase 2). Furthermore, as indicated earlier, executive decisions, which require autonoetic processing, or self-conscious distancing from one's present self, also provide the ground for developing a conception of an extended self. I believe that the extended self is also required for, or develops concurrently with, the development of a theory of mind (again see Table 8.3, Phase 2). The evidence that older 3-year-olds are just acquiring the ability to do the delayed self-recognition task, and that this correlates with executive function and with the choice task, suggests that it is during this period that all these activities are going through concurrent development. By contrast, the third experiment seems to indicate that the primary effect of executive function on choice activity has already occurred, and from the second experiment, it appears that the primary effect of theory of mind has yet to appear. Thus it seems most reasonable to suppose that it is out of executive decisions involving a future self that children first come to think of themselves as extended in time. Furthermore, it is out of this forward-looking decision making activity that children eventually develop an abstract conception, or theory, of a temporally extended mind that can be applied to others as well as to the self.

HAZLITT AND THE EXTENDED SELF

Hazlitt (1805/1969) made two important contributions toward understanding the child's development of the concept of an extended

self. The first was his realization that the only way for one to connect to the future was through imagination—what today is called *simulation*—and that the use of the imagination to represent future selves applies equally well to the future selves of other people as to one's own future self. Thus, insofar as sympathy with a future self involves the capacity to represent that self, and to identify with it, both future-oriented prudence and future-oriented altruism should, at least initially, be on par with each other. He realized that there is no necessary preference for one's own future over that of another person, because both depend on an imagination and sympathy—or identification—that takes one out of one's current point of view so that one can adopt the point of view of another temporally distinct self. Research conducted on the basis of this insight of Hazlitt provides some support for this view of the concurrent development of future-oriented prudence and altruism.

Hazlitt's second contribution was his suggestion that one must learn to give preference to one's own future self over that of another and that this learning involves treating one's future self in a way that one does not treat the future self of another. Thus far, I have not specifically considered this aspect of Hazlitt's view, but it is a significant one for appreciating the transformation that occurs in the child's conception of self during the age-4 transition. What develops at this time is a way of thinking about the future—and also about the past—that treats temporally distinct selves in a new way. These selves, which one can imagine, and remember, are partitioned into those that belong together and those that do not. The ones that belong together are those that form phases in the life of a single, continuing, thinking being, a self-conscious being with a past, present, and future, a being with a sense of personal identity through time. The self of the present comes to appropriate as his or her own those past and future states that are viewed as connected through consciousness to the present self and at the same time reject as other those past, present, and future states connected through consciousness to another person. Whereas a younger child knew only of a distinction between the present self and present other, and might be able to imagine other "present" selves distributed in space and time, the child now conceives of the self not as present but as extended. In doing so, he or she not only makes stronger connections to his or her own past and future but also builds a new barrier between one person and another, between a self as temporally extended and other temporally extended selves. Whether this new barrier should be seen as an unfortunate one because it facilitates self-interested thinking over altruism, as Hazlitt thought, or instead be seen as a fortunate one because it extends the range of activities of an integrated self, including moral activities, is one that cannot be answered here. In any event, the acquisition of the concept of a temporally extended self is an important event in human development, one with enormous implications for the

type of existence we have as a species. For it is through this concept of a temporally extended self that one can rise above the present self and can live in an abstract realm not localized in space and time as a rational being in communion with other rational beings.

ACKNOWLEDGMENTS

This chapter was prepared with support from the Social Sciences and Humanities Research Council of Canada (Grant 410-97-0767). I thank Chris Moore for very helpful comments on earlier versions of this chapter.

REFERENCES

Barresi, J. (2000) Intentional relations and divergent perspectives in social understanding. *Arob@se: Journal des Lettres et Sciences Humaines*, *4*, 74–99.

Barresi, J., & Moore, C. (1993) Sharing a perspective precedes the understanding of that perspective. *Behavioral and Brain Sciences*, *16*, 513–514.

Barresi, J., & Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences*, *19*, 107–122.

Carlson, S. M. (1997, April). *Individual differences in inhibitory control and children's theory of mind*. Poster presented at biennial meeting of Society for Research in Child Development, Washington, DC.

Descartes, R. (1984). Mediations on first philosophy. In J. Cottingham, R. Stoothoff, & D. Murdoch, (Trans.), *The philosophical writings of Descartes*, II (pp. 1–62). Cambridge, England: Cambridge University Press. (Original work published 1641)

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

Harris, P. (1991). The work of the imagination. In A. Whiten (Ed.), *Natural theories of mind* (pp. 283–304). Cambridge, MA: Basil Blackwell.

Hazlitt, W. (1969). *An essay on the principles of human action and some remarks on the systems of Hartley and Helvetius*. Gainesville, FL: Scholars' Facsimiles & Reprints. (Original work published 1805)

Hoffman, M. L. (1976). Empathy, role-taking, guilt, and development of altruistic motives. In T. Lickona (Ed.), *Moral development and behavior* (pp. 124–143). New York: Holt, Rinehart, and Winston.

Hughes, C. (1998a). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, *16*, 233–253.

Hughes, C. (1998b). Finding your marbles: Does preschoolers' strategic behavior predict later understanding of mind? *Developmental Psychology*, *34*, 1326–1339.

Hume, D. (1888*). Treatise of human nature* (L. A. Selby-Bigge, Ed.). Oxford, England: Clarendon. (Original work published 1739)

Locke, J. (1975) *An essay concerning human understanding* (P. H. Nidditch, Ed.). Oxford, England: Clarendon. (Original work published 1694)

Martin, R., & Barresi, J. (1995). *Hazlitt on the future of the self*. Journal of the History of Ideas, 56, 463–481.

Martin, R. & Barresi, J. (2000). *Naturalization of the soul: self and personal identity in the eighteenth century*. London: Routledge.

Mischel, W. (1974). Processes in delay of gratification. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 7, pp. 249–292). New York: Academic.

Moore, C., Barresi, J., & Thompson, C. (1998). The cognitive basis of future-oriented prosocial behavior. *Social Development*, *7*, 198–218.

Nagel, T. (1970). *The Possibility of Altruism*. Oxford, England: Clarendon.

Neisser, U. (1988). Five kinds of self. *Philosophical Psychology*, *1*, 35–59.

Nelson, K. (1992). Emergence of autobiographical memory at age 4. *Human Development*, *35*, 172–177.

Perner, J. (1992). Grasping the concept of representation: Its impact on 4-year-olds' theory of mind and beyond. *Human Development*, *35*, 146–155.

Perner, J., & Ruffman, T. (1995). Episodic memory and autonoetic consciousness: Developmental evidence and a theory of childhood amnesia. *Journal of Experimental Child Psychology*, *59*, 516–548.

Povinelli, D. J. (1995). The unduplicated self. In P. Rochat (Ed.), *The self in infancy: Theory and Research* (pp. 161–192). Amsterdam: Elsevier.

Reed, E. S. (1994). Perception is to self as memory is to selves. In U. Neisser & R. Fivush (Eds.), *The remembering self* (pp. 278–292). Cambridge, England: Cambridge University Press.

Russell, J., Jarrold, C., & Potel, D. (1994). What makes strategic deception difficult for children—the deception or the strategy? *British Journal of Developmental Psychology*, *12*, 301–314.

Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The "windows task" as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology*, *9*, 331–349.

Thompson, C., Barresi, J., & Moore, C. (1997). The development of future-oriented prudence and altruism in preschoolers. *Cognitive Development*, *12*, 199–212.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *25*, 1–12.

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin*, *121*, 331–354.