On becoming a person

JOHN BARRESI

ABSTRACT How does an entity become a person? Forty years ago Carl Rogers answered this question by suggesting that human beings become persons through a process of personal growth and self-discovery. In the present paper I provide six different answers to this question, which form a hierarchy of empirical projects and associated criteria that can be used to understand human personhood. They are: (1) persons are constructed out of natural but organic materials; (2) persons emerge as a form of adaptation through the process of evolution; (3) persons develop ontogenetically; (4) persons are created through the unifying activity of self-narrative; (5) persons are constituted through socio-historical and cultural processes; and (6) the concept of person is a normative ideal. I suggest that it is important to consider all of these projects and related criteria in order to appreciate fully how an entity becomes a human person.

About 30-years ago I first read Carl Rogers’ book, On becoming a person (1961). It was a revelation. At the time I knew little of psychology, but I was concerned with what to do with my life, where to find meaning. Rogers dealt with these concerns. At the same time he provided insight into how to understand scientifically what it is to become a person. So I got hooked on psychology, little understanding how unusual this concern of Rogers was for the profession.

Central to Rogers’ book was the chapter titled “Persons or science: a philosophical question” previously published as an article in the American Psychologist (1955). It was a working paper, in which he considered in two ways his relationship to his clients: as a practicing clinician and as a scientist. As a practicing clinician he felt that behaving naturally seemed to help his clients more than acting as a scientist with specialist knowledge. Yet he felt he could understand the process of change in his patients best by donning a scientific hat. He revealed the conflicts that arose from these two roles, and tried to resolve them. One suspects that Rogers never did resolve them, since he eventually moved out of academic psychology to become an applied psychologist.

Although it took me a number of years after my own training as an academic psychologist to return to this issue, I maintained a continuing interest in what is it to become a person. My interest became a desire to see the development of a science of person, a personology. This would not be a science of personality, focusing on

John Barresi, Department of Psychology, Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada; e-mail: JBARRESI@is.dal.ca

0951-5089/99/010079-20 $7.00 © 1999 Taylor & Francis Ltd
types of persons or on general principles of the development of personality; nor would it be a clinical and applied psychology of the person, focusing on abnormalities of personality and the means to provide cures or adjustments. Instead the focus would be on: (1) what a person is; (2) what it takes to make an object or organism into a person; and (3) how it becomes one. The present paper—like Rogers’ original paper—is a progress report on my attempts to deal with these questions, but it is also a report on what I see as an emerging general interest in personology.

Defining the concept of person

When we think of what a person is we usually think, quite simply, of a human being. However, person is a psychological category, while human being is a biological one. Historically, the term person was defined as “an individual substance of a rational nature” (Boethius, circa fifth century AD), a definition based on Aristotle’s thought, and applied generally to human beings as well as to immaterial spirits, such as angels and God. In the case of human beings, our rational essence was associated with the concept of an immortal and immaterial soul. This “spiritual” approach to the concept of person carried itself even into the origins of scientific psychology in the 17th and 18th centuries. Only with the development of experimental and materialist approaches to the science of the mind has the category of person collapsed into that of human persons (see Martin & Barresi, 1999).

Even so, the individuality of self or person is not tied to the mere biological identity of a human individual, but rather to the psychological identity of self. It was Locke’s (1694/1975) definition of person in terms of continuity of self-consciousness that both maintained the distinction between person and human as categories, and provided an empirical framework within which to develop a science of persons. In making this historic contribution, Locke defines person as:

a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing in different times and places; which it does only by that consciousness which is inseparable from thinking, and, as it seems to me, essential to it. ... For since consciousness always accompanies thinking, and it is that which 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 it is by the same self with this present one that now reflects on it, that action was done. (Locke, 1694/1975, p. 335)

Locke thus went beyond the classical definition of person as an individual rational substance or intelligent being. Instead, following in a tradition begun by Descartes, he focused on the capacity of a rational being to be conscious of itself in the very acts of thinking and perceiving, to be aware of itself as the subject of thoughts and
perceptions, and to recognize itself as the same self that had other experiences at other times and places.

As we now know from cases of multiple personality, it is possible for a single human being to have several apparently independent self-consciousnesses, each with its own store of past memories of its own activities. (Locke actually discussed the possibility of multiple persons attached to the same human body, but—so far as we can tell—merely as hypothetical possibilities.) Moreover, we have reason to believe that consciousness of oneself as the constant source or subject of one’s activities, one’s thoughts, perceptions, and memories is not a capacity that exists for all humans. For instance, empirical evidence suggests that autobiographical memory in children, insofar as it deals with their previous thoughts and perceptions, first occurs in four-year-olds, at a time when they develop what is called a theory of mind. Hence, while Locke’s definition of person certainly captures a concept that often applies to normal adult human beings, it does not apply universally and straightforwardly to all human beings. And it is at least conceivable that his definition might be true of other organisms and possibly inanimate objects, such as robots, that are not humans. As such, it might be better understood as a concept providing criteria of personhood that can be applied to any being or object, than as one that applies only to human beings.

Criteria of personhood

Rather than attempting a definitional approach to the concept of person, such as Locke’s, it may be more fruitful to try to develop criteria by which to distinguish persons from other animate and inanimate objects. There is reason to believe that Locke himself may have thought of personhood in terms of multiple criteria. After giving his definition of person, he states:

*Person*, as I take it, is the name for the *self*. Where-ever a man finds, what he calls *himself*, there I think another may say is the same *Person*. It is a Forensick Term appropriating Actions and their Merit; and so belongs to intelligent beings capable of Law, and Happiness and Misery. This personality extends it *self* beyond present Existence to what is past, only by consciousness, whereby it becomes concerned and accountable, owns and imputes to it *self* past Actions, just on the same ground, and for the same reason, that it does in the present. (Locke, 1694/1975, p. 346)

Here, Locke is concerned more with a criterion for moral than for metaphysical personhood. As Dennett (1978) points out, these two criteria need not pick out the same objects. For instance, it is possible that a six-year-old child might satisfy the criterion of self-consciousness, yet not be capable of moral responsibility for her/his actions.

Dennett (1978) considers six conditions for personhood: (1) persons are *rational beings*; (2) they are beings or objects to which *mental* (intentional) *predicates* or *states of consciousness* apply; (3) we take a *personal* stance toward them; (4) they are capable of taking a *reciprocal* personal stance toward us; (5) they are capable of *verbal*
communication; and (6) they are capable of a form of self-consciousness not found in other species. Dennett sees these criteria as forming a hierarchy, where the lower ones must obtain before the higher.

In what follows, I formulate a framework for investigating human personhood that like Dennett’s criteria, yields a hierarchy, placing increasingly stringent demands on what is necessary to become a person. But my goal is different than Dennett’s. It is not to defend a philosophical account of personhood intended to be applicable to any sort of potential person (e.g. robots) and to be used for such projects as normative ethical theorizing. Rather, my goal is to present a nested set of empirical projects and empirically derived criteria that contribute toward a richer understanding of human personhood and might be useful in guiding further research into human personhood. I do not deny that there may be other kinds of persons, the projects and criteria for which might follow a different track. Nor do I claim that only the levels of personhood that I consider are important to human personhood.

Each of the empirical projects and associated criteria that I suggest will restrict the range of objects that can be considered persons. The first project and criterion will distinguish inanimate from animate objects, and exclude as persons, as Dennett’s criteria would not, even robots that appear to be functionally equivalent to human persons. Until we understand better the material ground for our own personhood, it seems to me to be an unwarranted assumption to suppose that any functional criterion of personhood that we might try to use, that excluded our material base, will succeed in capturing all that is essential to human personhood. The second project will investigate those factors in evolutionary history that were necessary for making possible a species in which individuals are capable of becoming persons. As we shall see, one of the outcomes of this investigation of evolutionary origins of human personhood is the hypothesis that certain forms of self-consciousness of a human kind may have preceded the evolution of verbal language. If this is confirmed in further research, it would argue against Dennett’s suggestion that verbal language is necessary for self-consciousness. The third project looks into empirical discoveries regarding an important early phase in human development when children acquire an initial capacity to become conscious of themselves and others as mental agents or selves extended in time. It is this capacity for an extended self which seems to separate human from other primate species, and thus may provide a crucial criterion for human personhood. A fourth project and criterion emerges in the investigation of late adolescent children and adults. It focuses on our ability to create an identity or self through narrative. It is at this point that Dennett’s language requirement, may, indeed, play a crucial role, though not in isolation but in conjunction with other capacities that occur in late adolescence and that have been extensively investigated. I suggest that this capacity may be required for an individual to become a full moral agent. A fifth empirical project and criterion is concerned with variations in the kind of identity allowed humans within a society or culture, as a function of historical developments. This investigation suggests that without the support of one’s culture one cannot achieve full personhood. Thus, social opportunity must be added to personal capacity to achieve the highest levels
of personhood. Finally, a sixth project and criterion considers personhood as normative ideal, rather than as a currently existing property or capacity.

The construction of persons

On 10 February 1996, during the first game of his initial match with Deep Blue, Garry Kasparov, the leading human chess champion, faced a move by the leading inanimate chess champion that seemed wrong. It was too human a move for a calculating computer to choose. Kasparov sensed that he was facing a “new kind of intelligence” (Kasparov, 1996). Deep Blue foresaw consequences that Kasparov could not. And so in this first game of the championship, an inanimate object “out-thought” humanity’s best “thinking thing.” However, the computer’s advantage was only temporary. Ultimately, Kasparov was able to defeat the machine by making some trivial, actually rather meaningless, moves, which disoriented the machine. Because these moves broke patterned structures that the machine would normally recognize from its large store of past games, the machine had to use a weaker strategy of calculating all possible moves into the future. Since even this machine could not yet calculate all these moves within its time limit, it failed to succeed where other human players, who could recognize the deceptions that Kasparov was engaging in, might have succeeded. In short, Kasparov avoided situations for which the program’s calculation possibilities would give it an advantage by making disorienting moves which led to situations in which he, with his human intuition, would have the advantage.

During this first encounter with Deep Blue, Kasparov pitted human flexibility against Deep Blue’s computational advantages. But in the second match, a year later, Kasparov ultimately lost to Deep Blue. In losing, what most troubled Kasparov was the apparent flexibility of Deep Blue. Again the machine sometimes made moves that seemed much too human for a computer. Only this time, these moves were not explicable in terms of long-range calculations. Kasparov (1997) discovered that he was playing “a very flexible, quickly changing opponent” in addition to one “with an ability to avoid any mistakes in long term calculations” (p. 38). And Kasparov was deeply troubled and suspicious about this flexibility. After the match he wrote: “I believe the IBM team owes the world of chess, and the world of science, a full explanation of how such a flexible machine was developed. They have to make all the scientific data available to allow others to judge their accomplishment” (p. 39). Kasparov wanted a rematch and planned to change his own strategy. He would not play under the assumption that his opponent was a machine. “I think this match proved that there should be no special anticomputer strategy. To beat this machine, I just have to play great chess. ... I need physical and psychological stability, a great level of concentration and a mind free of other distractions to calculate, calculate and calculate” (p. 39).

Despite all the hype about this confrontation between animate and inanimate “thinking things,” Deep Blue is still only a machine, but one with enormous computational power and sophisticated software. Its behavior is convincingly intelligent, so that even the greatest human chess player is at a loss on how to take
advantage of its inanimate nature, and must instead only hope to match it in calculating power. Yet, by no means would we be inclined to mistake Deep Blue for a thinking being—or person. Deep Blue satisfies the first three of Dennett’s criteria of personhood: it behaves rationally and its behavior can be interpreted in terms of goals, purposes, and other mental (intentional) states. Kasparov certainly feels that at least with respect to its ability to play chess, it should be treated as a person, not an inflexible machine. But it is also certain that Deep Blue cannot reciprocate this personal stance. It has no knowledge of itself or others as rational agents with purposes, nor does it have any concept of what a person is, or how to take a personal stance with respect to others. Furthermore, it does not exhibit linguistic abilities, or any form of self-consciousness.

Still, Deep Blue is a product of only the first 50 years of computer technology and of our increasing understanding of our selves (metaphorically) as information processing machines. What will the next century bring? How far will we advance in neuroscience and cognitive science toward an understanding of our own intelligence, and our capacities for personhood? Will we be able to create machines that satisfy the criteria of personhood that we apply to ourselves?

Currently much of the debate is over whether we can ever overcome the “hard” problem of consciousness (Chalmers, 1995, 1996). We can imagine someday solving the “easy” problem of creating robots that are our functional equivalents—robots that “process information” and generate behavior the way we do. Advances in neuroscience, cognitive science, and artificial intelligence are so quickly discovering how minds/brains work that many anticipate making functional equivalents to ourselves in the not too distant future. But will these robots be persons, or “zombies”—where a zombie is a machine that acts like a person but lacks consciousness? Some people, like Dennett (1991), argue that zombies are conceptual impossibilities, because once we create functional equivalents to ourselves, then they will be as conscious as we are. Dennett believes that consciousness itself can be explained in terms of the functional operation of our cognitive/behavioral systems. There is no need to postulate any extra “qualia” or experience that cannot be captured through the construction of information processing systems that do what we do. In contrast, Chalmers argues that conscious experience is something extra that cannot be reduced purely to information processing, though he believes that it is strongly correlated with some such activity in human beings. The “hard” problem is to figure out a coherent theory of consciousness, on the assumption that it is scientifically possible to create zombies that do what we do but do not experience what we experience. Why are humans conscious, if we can conceivably create functional equivalents that are zombies? And how does it happen that physical systems produce experience? These are the hard questions for which we have as yet no answers.

The real issue here is whether consciousness has functions relevant to human personhood that cannot be captured in an unconscious robot, or zombie (cf. Barresi, 1987, 1995). If consciousness does, we will have to understand how it is physically possible and functionally relevant to our activities as persons in a way that excludes the possibility of creating purely functional equivalents that are themselves
unconscious. Otherwise, we will have to understand the functions of conscious experience in a way that can be duplicated in unconscious robots, which will require a much clearer understanding of these functions and how they relate to our personhood than we currently possess.

For now, we have reason to believe that the mere appearance of behavioral equivalence between a robot and a human person cannot be all there is to personhood in human beings. In a previous publication (Barresi, 1987), I have argued against the possibility that any Turing-like test will reveal all the relevant functions of human personhood. All that is required for a machine to pass a Turing test is to convince other human beings that the machine functions as a person. But to convince other humans of this and to actually function as a person, all the way down, are entirely different things. For instance, a Turing-like test may fail to reveal functions relevant to our personhood that have accumulated through evolution, such as certain internal creative capacities associated with consciousness, or certain intellectual or emotional capacities necessary for our social adaptation, but which we do not perceive. So, mere functional descriptions that we can come up with based on behavioral equivalence may not fully reveal crucial functions that make human personhood possible.

More is required. We must also understand how human personhood is—in fact—constituted. For instance, if we understand human personhood as a material process, then we must understand its organic base, and how the base produces whatever sorts of conscious experience that we view as essential to personhood. Currently, we have no idea how such a transformation of matter is possible. No current conceptions of information-processing in inanimate machines even begin to bridge this theoretical gap. What we do know, based on the continuity between humans and other organisms, is that there must have been a long line of evolutionary development that led to the emergence of beings that are not only conscious but also self-conscious in the way that is required to become a human person. Thus, I will assume that persons are necessarily animate machines or organic beings, and try to determine what it is about their evolutionary origins and natural development that makes it possible for them to become persons.

The evolution of persons

It may be that we can never succeed in building a robot or cybernetic man that is also a person, for we can never understand what it is that makes us persons (Barresi, 1987). But if the theory of evolution is right, then certainly it must be possible to make persons, since nature has done it; otherwise we wouldn't be here. Maybe a better way to try to understand what it is to become a person is to track nature's way of building persons. One advantage of taking an evolutionary approach is that we can expect it to yield an answer to the question: "Why are there persons?"

Trying to reconstruct an evolutionary history of our species is a daunting task, since we have so little information about early hominids and their adaptations. Hence, I will not try to provide a detailed phylogenetic reconstruction, but instead will merely recommend one (Donald, 1991) that seems particularly relevant to the
issues of personhood. The main fact that has to be dealt with is that particularly with respect to brain size, later hominid species split off from other primates and early hominids. From the earliest hominids to current humans, there was a threefold increase in brain/body proportion. It seems likely that this evolution in brain size—particularly with respect to the neocortex—is associated with a form of adaptation involving an intense and complex social life in which increasing need for cooperation within groups and competition between groups (as well as between related species) was causally effective in producing the rapid evolution of brain size. Furthermore, it seems likely that among the earlier capacities to emerge was the capacity for spontaneous imitation and the ability to represent action sequences extended in time and directed at future events, and that long before the development of a verbal symbolic and arbitrary representation through narrative, the symbolic representation and communication of such activities were initially through mimicry. Given these plausible elements of our evolutionary history, we can look more closely at how human beings compare to their closest primate relatives, in order to evaluate further how capacities and behaviors associated with various criteria of personhood may have emerged through evolution.

Recent evidence suggests that the capacity for spontaneous and accurate imitation of actions and mental attitudes is a capacity essential for the emergence of culture and of the appreciation of mental states in other individuals (see e.g. Tomasello et al., 1993; Barresi & Moore, 1993, 1996). Barresi and Moore have argued that imitation is also essential for acquiring a representational understanding of one’s own mind. Evidence for the hypothesis of a strong relation between spontaneous and accurate imitation and social understanding of self and other is indicated by the co-occurrence of both these abilities in only a few species. With the possible exception of cetaceans (e.g. dolphins) and some birds, there is little evidence, other than in the higher primates, of spontaneous imitation in organisms. Furthermore, in chimpanzees (and possibly orangutans), where we see at least some evidence of such imitative capacities, we also have evidence that they have a kind of understanding of the intentional activities of both self and other that exceeds the capacity of lower organisms, such as monkeys.

According to Barresi and Moore, the essential difference is that higher primates and humans are able to represent the intentional activities of self and other within a single representational system, so that they can understand the activities of self and other in the same way. Lower organisms are hypothesized to understand the activities of self and other using different forms of representation. As a result, such organisms cannot learn the meaning of another’s activity from the other’s first person perspective by directly observing that activity; nor can they directly imitate useful goal directed activity that they observe. Furthermore, because they cannot understand their own activities from a third person perspective, they fail to have sufficient self-consciousness to recognize themselves in mirrors, which has been shown to be a distinctive capacity of higher primates (Gallup, 1970). But Barresi and Moore also note that as compared to humans, even chimpanzees are limited in social understanding because they cannot represent mental states directed at absent or counterfactual objects. Their capacity for social understanding is limited to the
here and now, and not to past and future mental states, or current mental states directed at non-existent objects. Only humans have such capacities. In other words, only humans have what is called a representational theory of mind (Perner, 1991).

We can now consider the phylogenetic implications for criteria of personhood of these observations. Of particular interest here is to see where the line has to be drawn between humans and other organisms on criteria relevant to personhood, and how well other organisms satisfy less stringent criteria. If we consider the criteria suggested by Dennett (1978), it is clear that only higher primates and humans (in the primate line) show evidence of satisfying his first four criteria. Although it is possible to ascribe (1) rationality and (2) mental states to lower organisms, it is not necessary to suppose that they make such intentional attributions about each other. It seems more reasonable, based on current evidence, to suppose that lower organisms take a causal or behaviorist attitude toward each other’s activities. By contrast, higher primates and humans show evidence of understanding each other’s activities intentionally, and of reciprocating intentional or personal stances toward each other (3 and 4). On the fifth criterion of language, only humans participate in a communicative symbolic language (unless one wishes to include language-using chimpanzees, like Kanzi). However, it may be that Dennett’s framework misses some meaningful relationships between criteria by making language a necessary condition for the sixth criterion of self-consciousness (cf. Rovanne, 1994). Evidence of a symmetry in understanding of self and other in higher primates and humans suggests that there may be levels of self-consciousness attributable to higher primates, though these levels are not as rich and general as linguistic—representational—forms of self-consciousness found only at the human level. Moreover, it appears that the evolution of verbal symbolic language in hominids was probably a late development, after most of the change in brain size and complexities in social structure had occurred. Hence, there is reason to suspect that early hominids achieved levels of self-consciousness quite similar in some respects to our own, though without verbal language.

Thus, focusing only on Dennett’s criteria, and on the conceptual possibility that some forms of self-consciousness very similar to our own can exist that do not require verbal language, it seems reasonable to suppose that a species with self-conscious persons probably existed at one time in the hominid line prior to the emergence of verbal language. This would indicate that verbal language is not, after all, as important a condition for personhood as Dennett thinks it is. Furthermore, when we consider existing organisms and their relationship to Dennett’s criteria, it appears that chimpanzees (and probably orangutans) not only surpass inanimate objects like “Deep Blue,” in Dennett’s criteria of personhood, but also appear to be self-conscious. These organisms not only exhibit intentionality (or, based on their behavior, can be ascribed intentionality), but also recognize each other’s intentionality and seem capable of maintaining “personal” relations with each other by representing some of both their own and another’s current mental states. Furthermore, the fact that they may be able to represent some of their own mental states and are able to recognize themselves in a mirror is at least presumptive evidence that they are self-conscious.
Yet, even putting aside the role that language may play in human consciousness, chimp self-consciousness is distinctive from our own. The chimpanzee’s understanding of mental states of itself (and others) is restricted to the here and now, and, hence, it does not have any autobiographical understanding of its past self; nor can it anticipate its own future self as humans can. So, though chimpanzees have an understanding of themselves (and others) in the present, they do not have consciousness of what Neisser (1988) has called an “extended self,” that kind of self-consciousness that was Locke’s concern when he defined personal identity in terms of a temporally extended self-consciousness. It is this capacity to represent our selves and others as extended in time that primarily distinguishes hominid from primate possibilities for personhood, so it is in this difference that we must search for a function that distinguishes hominid from primate personhood. One possibility here is that the emergence of a cultural or moral order depends on one’s capacity to exist as an extended self, with moral responsibilities that extend through time, and that the main difference between primates and hominids is in having this ability to act as an extended self toward other extended selves. It may be that it is this difference which separates a merely protocultural species like chimpanzees, whose personal relations exist primarily in the present, and a truly cultural species like humans, whose personal relations extend through time.

The development of persons

Although it is generally considered important to treat all human beings as persons, not all humans have the capacity to act as persons in the technical sense that we have been exploring. This is especially true of young children, whose capacities for acting as persons are extremely limited, especially with respect to those higher criteria of self-consciousness and moral responsibility normally ascribed to adult humans. Thus, there are grounds for looking into the ontogenesis or development of personhood. While we will focus in the next section on the late adolescent development of personhood as involving self-creation through narrative, in the present section we will focus on an earlier phase which shows young children beginning to surpass their primate relatives in their capacities to achieve personhood.

In the previous section, we saw that a distinction occurs between the chimpanzee and human being in their relative ability to achieve an extended self. There are capacities that the child achieves around the age of four years, all of which apparently surpass the capacities of chimpanzees and which seem related to the early emergence of the human form of personhood. Although there are a number of precursors that already distinguish us from chimpanzees, such as our very early capacity for engaging in shared attentional activity, and our use of a symbolic language, what develops at this time is a sophisticated understanding of the nature of mental phenomena that can be applied to both self and other. This “theory of mind” gives the child a form of self- and other-consciousness that is absent in the chimpanzee.

At this time the child acquires the ability to represent not only what someone feels and knows (i.e. perceives) about a current situation, as well as what a person
is doing, all of which seem to be within the capacity of the chimpanzee, but also intentional activities directed at non-current situations or objects. Thus, the child now can understand that a person who was absent while some change occurred in the situation may have a false belief about an object that was moved in his or her absence (Wimmer & Perner, 1983). They can also remember their own previous beliefs and desires, even though these previous mental states were directed at objects no longer in existence (see e.g. Gopnik, 1993; Moore et al., 1995). Furthermore, at this time the child becomes capable of representing the divergence in current point of view between her- or himself and another (Gopnik, 1993) and develops an “autobiographical self” (Nelson, 1992). It has been suggested that these phenomena are interrelated (e.g. Perner, 1992). One important piece of evidence in this respect is the recent work of Povinelli (1995) on the extended self, who has shown that children around this age first acquire the ability to use a videotape of someone putting a sticker on the child’s head a few minutes earlier in the session in order to realize, when the tape is replayed, that the sticker might still be there and so to remove it. This form of recognition of a self that is extended in time is not shown in three-year-olds, though they can certainly recognize that the video is of themselves, just as they can recognize themselves in the mirror from 18 months. Whereas 18-month-olds, as well as chimpanzees, can recognize current reflections of themselves, these new findings suggest a capacity for an integrated understanding of oneself and others as beings that extend and transform through time, a capacity apparently outside the range of chimpanzees, though they have not yet been tested on such a videotape task.

Thompson et al., (1997) have found evidence that the child’s capacity to deal with the future as well as the past advances at this same age. Children were asked to make choice decisions involving play stickers where the stickers could be received in the future or in the present, and either were received only by the child or involved sharing stickers with another person. Although even three-year-olds were more willing that another person also get a sticker than that they alone should get one in the present situation, only four- and five-year-old children were capable of exhibiting future-oriented prudence and altruism. Only the older children would forgo a single sticker immediately received for two stickers later for themselves, or one each later for themselves and the other person. Thompson et al., (1997) hypothesized that the reason younger children were unable to defer reward was due in part to their inability to represent the future mental states of self and other when these conflict with their own current states. Further research, in which theory of mind tasks were correlated with these choice tasks, indicated some evidence for such a relationship (Moore et al., 1998). There was also evidence in this study that the emergence of the “executive” ability to inhibit a potent response to a desired stimulus might be involved. This latter result can be interpreted as indicating that the first emergence of a “will” that differs from mere desire may be involved in the child’s ability to delay gratification in these choice tasks.

Taken together, these results suggest that the first emergence of the uniquely human capacity for an extended self-consciousness occurs in the four-year-old, and that it is importantly related to the development of the ability to deal with one’s
activities from a rational perspective that extends through time, but also in relation to other extended selves. As such, it may form the basis of a moral self that not only concerns itself with its own future interests, but with commitments and duties to others within a moral social order. Hence, as Locke suspected, there is a connection here between being conscious of one’s self as extended in time and becoming a moral agent. Although the child at this age does not yet have a well-developed will, one which is “free,” in the sense of being able to choose or constitute itself (cf. Frankfurt, 1971), we have here what seems to be a necessary condition for such a will. That condition is the ability to conceive of one’s self as extended in time, so that one can make commitments to one’s self and to others that extend retrospectively from the past and prospectively into the future. In the next section we will consider later developments in such an extended self-consciousness more directly related to self-constitution.

Narrative creation of persons

One of the unique aspects of human self-consciousness or personhood is our ability to connect the distant past through the present and into the distant future through narrative, in particular through a narrative of our own lives. As many recent theorists have come to realize, it is through such story-making or story-telling that we actively create or constitute ourselves as selves, for it is through such narratives that we rationalize the meaning of our lives as well as dedicate ourselves to long-term goals (e.g. Barresi & Juckes, 1997; Bruner, 1990; Carr, 1986; Dennett, 1992; Hermans et al., 1992; McAdams, 1990; MacIntyre, 1981). As we have seen, the four-year-old child is already beginning to conceive of her/himself within a framework of an extended self—a self with a past, present, and future open to narrative connection. However, at this early age the child does not construct life-narratives. It is not until adolescence that humans begin to think abstractly about their lives and pass through the period that Erikson has called the “identity crisis” (Erikson, 1968), and construct the earliest narrative interpretations of their lives, which are then continuously revised and extended (McAdams, 1990). It is also at this time that individuals are faced with the possible social identities available to them within the culture. The self-identity that they construct for themselves must be congruent with the possible identities that are afforded them within the culture. Otherwise, their self-constituting acts, however satisfying as providing coherent narrative identities, will not insure social recognition of these identities, let alone personhood within their culture. We will consider this social determinant of identity and personhood in the next section. In the present section we will focus on the self-constituting aspect of narrative in the construction of identity.

Ever since Locke introduced his definition of personal identity that based it on self-consciousness, most Western thinkers have thought that self-continuity required some form of connection between stages of a self through time that depended on self-consciousness, and that this continuity of self-consciousness was constitutive of personal identity. Even Hume, in his “fictional” account of personal identity,
supposed that we form a natural belief in our own identity through imagination and memory. This self-consciousness appeared as an inner recognition, or belief, that it was one’s “self” who had performed the acts in the past that one remembers having done. And this same consciousness of a self allows us to look forward to the future, confident that it is us who will experience the consequences of our current acts, and remembering that we had been their cause. From a phenomenological perspective, it does not matter whether this self that we attribute identity to through time is metaphysically real or mere psychological fiction, what matters is that it is essential to our phenomenology of self, and without it, we scarcely could consider ourselves as persons, with a past, present, and future.

We have already seen how the notion of such an extended self that connects the past to future selves becomes a possibility in young children. What is absent from this view of an extended self based purely on self-consciousness is the active role that the individual takes in the constitution of the self. Locke himself introduced the relevant concept of “appropriation” to indicate the active role we take in self-constitution. He suggested that we “own” or appropriate our past acts as our own through re-collective self-consciousness and moral attributions. He assumed that re-collective consciousness automatically self-appropriated past acts, for which we accepted responsibility. Subsequent thinkers, however, have made it clear that such appropriation is not an automatic phenomenon, and that the individual is actively involved in accepting and rejecting—as its own—past and even present actions or behaviors of the individual.

William James (1890/1950) was one of the most important theorizers on consciousness and the self to focus special attention on the activity of appropriation in consciousness. He viewed each thought in the stream of consciousness as its own subject, or “I,” that would experience a current content of self, or “me,” then extend this content toward the past in a selective manner, appropriating some of the past but not all of it. (For a more recent conception of self-continuity through time based on appropriation, see Martin, 1998.)

The narrative approach to self recognizes the importance of the appropriative activity of consciousness in the constitution of the self, but it goes further. It also recognizes that the appropriative process typically involves a storyline, where the self is the chief character in the drama or plot of a life, that other people have secondary roles to play, and that the physical and social world is the stage upon which the leading character—the self—plays his or her part. What is essential to the formation of such a self is to have a stable identity through time, which one constitutes by acting within the play in which one conceives oneself to be engaged. We narrate the story of our lives to ourselves as we live it, and to others when they ask us who we are and what we are about. In narrating our life, we construct our own identities, to which we try to adhere until we are provoked to revise the story. However, because the narratives are generated at particular points in time and from particular and ever-changing authorial viewpoints, often the identities that we form in our self-narratives conflict with each other.

In multiple personality disorder, such conflicting selves take their most dramatic form. It will help clarify the viewpoint that we have been presenting here if we
consider in some detail the nature of the conflict between selves in a particular case of multiple personality (see also Humphrey & Dennett, 1991, for a related view). Each of us normally has one continuous connected personal consciousness upon which we construct narratives of what we are about. However, in the case of multiple personality, there can be several such streams of consciousness, often isolated from each other, with unaccountable breaks in time, when a given personal consciousness is merely latent or asleep and another active. There are also cases, such as Morton Prince’s famous case of multiple personality, B.C.A. (see, Barresi, 1994), where one personal consciousness is aware of some of the activities of a parallel personal consciousness, while maintaining her own separate stream of thoughts. While such examples of several streams of thought within a single human being are unusual, their structural properties as well as the causal history of such cases can help us clarify the distinction between “personal” narrative and organismic history.

The case of B.C.A. is especially revealing on this score, because the several personalities actively engaged in providing Prince with autobiographical accounts of their emergence as a multiple personality in letters and autobiographies (see Barresi, 1994; B.C.A., 1908, 1908/1909). Although B.C.A.’s dissociative capacity seems to have originated in a frightening incident in childhood, the major cause of the final fracture of personality was a long period of marriage where B.C.A. experienced a divided consciousness. During this period one part of her personality maintained the stance of a devoted wife, while another part wished to escape from the sexual responsibilities of marriage to a man she respected but did not love. During his final illness, which lasted several years, the inner conflict between these two parts, which were both part of a single stream of consciousness, became intense. After his death, for most of a year the “devoted wife” dominated the activities of B.C.A., while the second part disappeared. But when she was sent to a sanitarium by her doctor, the suppressed side of her self soon re-appeared, stronger than it had been, and finally emerged as a second personality when she was suddenly kissed by another patient at the sanitarium. These two sides of B.C.A. now alternated but were not yet amnestically isolated until later when, under the care of Prince, the second personality, B, appeared under hypnosis, after which the devoted wife, A, no longer had memory for B’s activities, tough B was co-conscious of A’s activities.

What is important to note is the way that A and B were able to distinguish between those thoughts and actions which they would appropriate to their own self-narrative and those they could not acknowledge as their “own,” even when they could remember having performed them. A (the devoted wife) would refer to the activities of the playful, childlike B, as acts done in a dream which she could not possibly understand as done by herself, though she retained memory for them during the early phase of personality dissociation before amnesia formed. B, likewise, had contempt for the activities and dark thoughts of the devoted widow, of which she was co-conscious. When B went on to try to write a reconstruction of her own history, she related the suppressed thoughts that occurred during B.C.A.’s married life as her own, though not yet thoughts of a separate self, which only formed later at the sanatorium. Still, she could appropriate these thoughts as her
own because of their content and viewpoint, which was continuous with the self that
did finally emerge.

While each normal person typically forms only one stable narrative self, we all
do selectively remember our past experiences and actions and appropriate some
more than others as central to ourselves, as we continuously constitute our selves as
relatively stable identities in time. In a case like that of B.C.A., several such stable
identities come to be formed, but the process of constituting a self-perspective
through narrative construction remains the same. What differs is the extent to which
those aspects of life history which are rejected come to form selves or identities of
their own, which alternate with other selves.

Just as consciousness of ourselves as extended in time is essential to our
becoming purposive agents with a will, our narrative construction of a psycho-social
identity is essential to our becoming recognized players in the social game of life.
But, as we shall see in the next section, without the support of the culture in which
we live, having a recognized narrative identity is not sufficient for full social status
as a person in one’s society and, hence, not sufficient for full personhood.

Socio-historical and cultural constitution of persons

The life history of the individual is necessarily intertwined with the cultural history
of the society in which the individual lives. Having the capacity for creation of an
identity through self-narrative, though necessary for personhood, is not, by itself,
sufficient to constitute an individual as a socially constituted person. Each culture,
and each historical period within a culture, sets the social framework in which
individuals are afforded whatever opportunities they have to become constituted and
recognized as persons. This framework can divide those individuals who are allowed
the status of persons from those who are not. Thus, in many traditional cultures
there is a distinction between individuals who have or can acquire “names,” where
those names entitle them to specific powers and responsibilities as persons, while
others—commonsers, women, children, and slaves—though afforded identities, have
no names or status as persons. Even in our own culture, women have only recently
acquired a variety of rights and privileges—traditionally available only to males—
that are essential for the full exercise of personhood within our society. Without such
rights and privileges, the individual, regardless of having some sense of personal
identity, lacks the power to act as a free agent within the society, thus to exploit
those personal powers necessary to achieve full moral responsibility toward either
self or other.

The non-biological and essentially cultural nature of human personhood is,
perhaps, best understood with respect to the example of slavery. How it is possible
to treat other human beings as property no different from horses, cattle, land, and
other goods, that can be bought and sold, now seems to many almost beyond
comprehension. Yet in the United States alone, laws were constituted and enacted
even up to 1860 that made it illegal for fugitive slaves to be allowed their freedom
even in non-slave states; and Northern ministers preached that former slaves should
voluntarily return to their master “for the common good.” In some of the slave states
it was also illegal for any individual of African descent to enter the state as a free person, regardless of whether that individual was a free citizen in another state or country. That a country founded on the principle that “all men (sic) are created equal” could have maintained a culture of slavery, where personhood was defined in terms of ownership or racial descent, can be explained only in terms of the power of culture to create its own social reality, however inconsistent that reality is to its own stated moral principles.

Cultural identity, like personal identity, has its base in narrative. And these two narratives are intertwined, each serving to help constitute and transform the other. Carr, in *Time, narrative, and history* (1986), argues that it is the temporal structure of experience that determines this intrinsic structure of both personal and collective narrative. Cultural groups and nations constitute their own identities through time through origin myths and cultural narratives extended through time. Their very existence as a group or nation depends on these self-constituting stories with a coherent and continuous point of view. The importance of narrative can also be seen in terms of groups within cultures or nations and the development of stories that extend beyond individuals but fall short of nations. This has much to do with the processes of historical change which involve the re-narrativizing of historical events into stories more congruent with changes in the perspectives of emerging groups within a culture. Often particular individuals and their self-narratives serve as provocative seeds for general social change, just as these self-narratives are themselves made possible by transformations in the narratives of groups within a culture and the culture at large.

However, we must not be misled into thinking that stories alone are the transforming factor in social change. As Juckes and Barresi (1993) point out, cultural change depends on both objective and subjective factors. The existent social structure and objective conditions are as important in determining change as the subjective interpretations of individuals and groups within a culture. The narratives we tell are not the only determinants of what is real, though they are often what guide our actions. As objective social conditions change, in part through such actions, then new narratives are formed which attempt to interpret and explain such objective changes. Hence, there is a dialogue or “duality” between the actions of individuals and groups seen “objectively” and “subjectively” which together provide the motor for social change and transformed conditions determining personhood within a culture.

Again, an example can help indicate how socio-historical change which affects personal and social identity emerges out of this duality of objective and subjective factors. In “Stories of Scottsboro” (Goodman, 1994), Goodman depicts an incident that occurred during the Depression, when nine black teenagers were seized from a train on which they were traveling through Alabama as hoboes in search of jobs. Some white male hoboes, whom they had fought and thrown off the train, had reported them. When the police had the train stopped, among the hoboes still on the train were two white women who accused the blacks of rape. The blacks claimed that they never even saw the women, and all the substantial evidence argued against the occurrence of any rapes (e.g. doctors’ reports on the women showed no
psychological or physical signs indicating rape). Nevertheless, within a month the blacks were sentenced to death for the rapes. Although none of them were ever executed, their repeated trials and convictions became a focal point for social transformation of black–white and North–South relations during the 1930s and 1950s. All sides had their own "stories of Scottsboro" to believe, to tell, as well as to guide their own actions. Goodman shows how all those stories had their own biases, agendas, but also narrative coherence, and how the people involved found it difficult to understand how others could possibly hold another story, and so they reacted vehemently in response to each other. Yet through the process of such reaction and dialogue, as well as the changing conditions in which they occurred, social change was enacted. In the penultimate chapter of the book, we see Clarence Norris—the Scottsboro boy, who had escaped probation in Alabama during the 1940s and ran off to the North—eventually being pardoned by Governor Wallace, and aggrandized by the black leadership in Alabama, during the changed social environment of the 1970s. Despite all the problems in race relations that remained, here was evidence of a change in perspective whereby whites and blacks of Alabama could acknowledge their common humanity and personhood in a manner that was impossible at the time of the initial Scottsboro trial.

What this example demonstrates is that personhood as a status within a society is not determined just based on our biological, or psychological capacity to have an extended self, and to create a narrative identity. While these capacities are necessary, and distinguish human individuals, who can and who cannot achieve higher levels of personhood within a society, they are not sufficient. Unless one has the opportunity to exercise the powers associated with these capacities, in order to act as a free moral agent among other moral agents, one does not yet achieve full personhood in the human sense. However, there is a sense in which none of us really fully achieves this ideal. All individuals in society only more or less exhibit the capacity to act as free moral agents with respect to each other, as we often fail to treat ourselves and each other as persons, by our own ideals. This variation suggests that personhood is, ultimately, more an ideal than an actual fact of human existence, as we shall see in the next section.

**Persons as nominative ideals**

Human beings live not only in a physical and social world, but also in an ideal world. Ideals regulate our behavior as much as physical or social causes. How is this possible? Can we understand a being with such capacities within a science of persons?

So far, we have restricted ourselves to empirical issues associated with becoming a person. We were interested in explanations of how objects become persons that are open to scientific or deterministic/causal accounts. We have considered five hierarchically related research projects and criteria associated with or discovered within these projects. The criteria we considered were scientifically interpretable as properties or capacities, or actions of objects, or of collections of objects. And our goal has been to identify particular empirically based criteria, which, taken together, form a
coherent hierarchy of essential elements of human personhood. While much more research needs to be done to develop a fuller understanding of the empirical elements of human personhood, for instance, into the development of human will beyond childhood, we have identified some important criteria that have differentiated stages of human personhood.

Thus far we have tried to avoid strictly philosophical and normative criteria in order to focus on human personhood as an empirical phenomenon. Even where normative issues might have been raised, such as in the previous section, we did not raise them until the very end of the section. The finding that in some societies at certain times in their history some individual human beings were "persons" by the operative criteria of that society and some were not was treated as a mere social or historical, rather than a conceptual, fact about personhood. And it was assumed to be open to empirical investigation to provide an account of how or why—in a causal sense—the criteria of social or legal personhood shifted through time in a particular society or culture.

Although we did not consider them at the time, a number of philosophical and normative issues about personhood could have been raised in previous sections. Is it right to treat robots, or animals, differently from humans just because they don't satisfy certain criteria of personhood that apply only to our species? What is the true or best metaphysical account of what it is to be a person? Are human infants and fetuses persons? Even if they are not persons by definition or rational criteria, do they still have any personal rights that are ignored by these approaches to their personhood? These are philosophical and normative questions for which a strictly empirical approach to criteria of personhood can have no answer. What then do these philosophical and normative issues imply about criteria for personhood? Can a purely scientific approach to personhood succeed? And if not satisfactory by itself, is it a worthwhile enterprise?

My own view is that our attempts to understand ourselves and others as persons must go beyond the limits of any possible science of persons (cf Barresi, 1987). At the level of understanding, as opposed to explanation, we must approach ourselves and others holistically. Our scientific activities, regardless of their objects, can only be one of our activities and can never fully determine our normative or ethical ideals. I take the view that "is" can never fully determine "ought," at least when it comes to personhood. We can never fully determine as scientists the "best" definition or criterion of personhood. All we can do is clarify empirical implications of any criteria we adopt.

In treating ourselves or others as persons I believe that we must adopt regulatory ideals not strictly reducible to empirical description or theoretical explanation. As such, certain definitions or criteria of personhood provide ideals rather than indicating a capacity or actuality that we have or necessarily can achieve (cf. Dennett, 1978, p. 285). For instance, one of the more important of our conceptions is that persons are moral agents with free wills, who are able to pursue ethical ideals. While such conceptions have their ground in empirically knowable human capacities and socio-cultural life, they go beyond any scientific concept of person. As normative ideals, they set agendas for our pursuit of a just society composed of morally
competent and responsible agents who treat each other as persons. Whether such an ideal can be achieved, and what limits we must be willing to accept in our pursuit of it, is something that can and should be affected by what we learn about ourselves through a science of persons. For instance, science can teach us about the limitations particular individuals or species have which prevent them from participating as persons with respect to particular ideals. But science cannot by itself decide what this ideal can and should be. Nor, however, should ideals be generated and used that ignore factual considerations about human personhood such as those that I have considered in the present paper. What is required is some measure of attention both to the facts of personhood, and to ideals of personhood that are not totally inconsistent with these facts, though not limited by them.

Acknowledgments

Support for this paper by grants from Dalhousie University’s Research Development Fund and the Social Science and Humanities Research Council of Canada is gratefully acknowledged. The author wishes to thank Tim Juckes and, especially, Ray Martin for their useful comments on previous versions of this manuscript. He also wishes to thank the editor and two anonymous reviewers for their suggestions.

References


