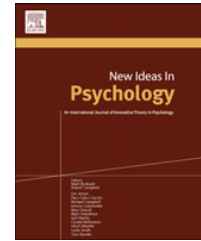




Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

New Ideas in Psychology

journal homepage: www.elsevier.com/locate/newideapsych



On seeing our selves and others as persons

John Barresi

Department of Psychology, Dalhousie University, Halifax, NS, Canada B3H 4J1

A B S T R A C T

Keywords:

Person
Self
Evolution
Reciprocal altruism
Extended self
Theory of mind
Moral psychology

Human beings may be the only organisms capable of thinking of self and other in equivalent ways – as selves and persons. Most organisms think about their own activities differently than they do the activities of others. A few large-brained organisms like chimps and dolphins sometimes think of the activities of self and other in the same way. But, only humans think quite generally in this manner. In this paper I give a description of our commonsense notions of self and person, and a scientific framework in which it can be fit. I then provide a phylo- and onto-genetic account of these concepts. Finally, I argue that the theory of reciprocal altruism provides the best account of why the notions of self and person evolved to have the form and function they do with respect to human social life and moral capacities.

© 2009 Elsevier Ltd. All rights reserved.

What are we? Among the obvious answers are that we are selves and persons. Each of us, including all of those who will read this paper, is a self; each is a person. But what is a self? A person? And are selves and persons the same or different things?

In current, everyday experience and language, being a self and being a person amount to the same thing. However, the use of one term or the other often depends on perspective. Person is a general term that we apply to any individual human; but self is a term that we use in a more restricted sense. Self in its everyday use is a term that indicates the special relation that holds between the person speaking or thinking and the person being referred to when they are the same person. In common experience, at least in English, 'self' and 'other' are opposing terms for the same kind of thing, an individual human being or person. They are deictic terms that shift with the user. I am a 'self' to me and an 'other' to you, and you are a 'self' to you and an 'other' to me. In order for us to use these terms we must be able to recognize that you and I are both persons and recognize our equivalence in this way. But our personhood is experienced differently by each of us so that your experience of your personhood is an

E-mail address: jbarresi@dal.ca

experience of your self and my experience of my personhood is an experience of my self, whereas our experience of each other's personhood is that of an-other person.

While everyday experience sees a close relation between self and person, theoretical use of these terms has followed common usage in some cases but not in others. This is true both currently and historically in philosophy, psychology, and other disciplines. For instance, Locke, in one of the most consequential discussions relating these terms (Locke, 1694), declared that: "PERSON, as I take it, is the name for ... self. Wherever a man finds what he calls himself, there, I think, another may say is the same person." This follows common usage in relating the two terms. But when he actually defines person and self, he gives them different definitions, neither of which matches our everyday meaning for these terms. He defines self as "that conscious thinking thing..., which is sensible, or conscious of pleasure and pain, capable of happiness or misery, and so is concerned for itself, as far as that consciousness extends." By contrast, his definition of person is "a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing in different times and places." While Locke's definition of self focuses on sensation and self-concern, his definition of person focuses on reflection and personal identity. Nevertheless, both of these definitions base identity of self and person over time on consciousness, which also connects the two: "[S]ince 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 that action was done" (Locke, 1694/1979, pp. 346; 341; 335). So, while Locke follows current, common usage in relating the two terms, he does not follow current, common usage in what he views as the self or person. It is not our embodied individuality, which makes us a person and self, but our capacity to think and experience our self now and as extended through time that makes us so, and this is based on consciousness rather than organic individuality.

In the present article, the focus will be less on philosophical, or other theoretical, accounts of what a person or self is, than on our commonsense or everyday notions of person and self and their function in human psychology. Historical changes in theory are discussed mainly for the insight they can provide on the status and transformations that these concepts have taken in Western thought. This will help to illuminate why the concept of person has lost its central status in psychology and why the concept of self has become diffused rather than occupying a central role in an integrated theory of person. Once we restore to theory our commonsense understanding of these terms, a more integrated approach to self and person and their roles in human psychology becomes apparent and once again available for theoretical development in some of the directions that I will outline in this article.

1. A brief history of theories of self and personal identity

As far as the history of selfhood and personhood goes (Barresi, 2006; Barresi & Martin, *in press*; Martin & Barresi, 2006), Locke's discussion was the major turning point between the early history, where person was a concept tied mainly to spiritual and religious notions of persons and selves, to more naturalistic notions that focused on consciousness and on our material nature. Before Locke, the most common definition took a person to be an individual substance of a spiritual nature. This definition applied to angels, humans, and God, but not to animals. In the opinion of most theorists, whereas persons had immaterial and immortal souls, animals had only material souls. The original Platonic notion of an immaterial soul that was the essential self had been naturalized in the late Middle Ages by Aristotelian philosophy with the notion that spiritual form and prime matter conjoined to form individual human, embodied persons. The soul on this view – at least as held by Aquinas – was not the self, but an incomplete part of a self or person. However, Descartes resuscitated and made even sharper the Platonic distinction between soul and body. For Descartes the self, or mind, is essentially a thinking thing and an individual immaterial soul, while the body is made of a distinct corporeal substance out of which all of nature, other than the human soul, was composed. Locke, who was influenced by Descartes but did not wholeheartedly accept his dualism, then proposed that consciousness and its interconnectivity be used in defining self and person, and left undecided whether this consciousness

was supported by a material or spiritual substantial base. For instance, he considered hypothetical examples where consciousness could transfer from one body to another, while maintaining personal identity. And he proposed that our resurrection after our bodily deaths might involve some process of this sort, whether or not immaterial souls are also involved.

As a result of Locke's influence, most subsequent philosophers as well as many psychologists have emphasized the role of consciousness in selfhood, and have either attached the concept of person to that of self as Locke did, or deflated the notion of individual embodied person relative to that of self. While giving up the notion of immaterial soul, except in religious thought, the notion of a unified mind combined with a changeable material body, continued to dominate theories of self well into the 20th century. As a result of these developments, theory moved away from our commonsense notion that self and person are terms that both refer to embodied individual objects but from different points of view, to a notion of self as unified conscious agent and person as an ill-defined combination of a conscious self with a body, or as with Locke, merely as another name for the conscious self.

Even so, beginning with Darwin and the rise of materialism in the 19th century, there has been an opposing empirical tradition that has focused on embodiment, though, in general, it has not given status to the notion of person. In psychology and in neuroscience, the emphasis has been on the material nature of the human organism and of mental phenomena as emergent properties of matter that depend heavily on brain activity that is either non-conscious or unconscious. As a result of the development of this tradition, the supposedly unified conscious self has become diffused into a variety of hyphenated self-related phenomena, where the self that is posited can be the body, the model of self in the brain, or various self-concepts acquired during development, but certainly not a unified subject of experience, as proposed in the main historical tradition that arose out of the notion of immaterial soul as self.

Joining with this biological orientation was a new social orientation to the self that began with Hegel, was reinforced by Marx, and was adopted in the social sciences, including social psychology, from the beginning of the 20th century. The self as an object of knowledge, rather than a subject of knowledge and an independent agent of activity, became a social concept of significant importance, one strongly dependent on socio-cultural and historical tradition. In the late 20th century this concept became even more diffused, as it became apparent that multiple social contexts implied multiple social selves constantly in flux, so that a single integrated notion of self as both subject and object of knowledge and action no longer seemed viable. Despite increasing interest in the phenomena of selfhood – the self was viewed only as an illusion held by human beings when thinking about their activities. As a result of this alternative tradition, not only was the commonsense notion of a person as a unified self-conscious organism lost, but also was the self as a single subject and object of consciousness lost.

There were, of course, some theorists who maintained a notion of the unity of self and person. Within psychology these theorists tended to be oriented toward understanding the life history of individual persons rather than viewing the individual as a collection of psychological processes (see Barresi & Jukes, 1997 for a historical review). Collectively, these theorists formed a subdisciplinary focus in the area of personality and called themselves personologists – a label conceived by Murray in his seminal book, *Explorations of Personality* (1938). Perhaps the most well known theorist who took this stance during the 20th century was Rogers (1955, 1961), for whom the concept of person was central. He approached the person from a first-person empathic view of one human person trying to understand another, as well as from a third-person point of view of a clinical scientist trying to understand psychological processes in his patients, and he postulated that a fully functioning person is an actualized self. More recently, McAdams (1990) has integrated Erikson's (1980) personality theory of stages in the life cycle with the role that self-narrative plays in achieving and maintaining self-identity for a person into a theory of person as life-story. While these various approaches to personology are fruitful for thinking about humans as persons, they do not satisfy all our needs for a psychology of personhood (Barresi, 1999). Some of what is left out of personology is an account of the evolution and function of personhood, as well as a developmental account of our commonsense understanding of our selves and others as persons. We also need a concept of person that can satisfy theoretical requirements of a variety of disciplinary and interdisciplinary contexts. It is toward the development of a general theoretical framework for the science of personhood of this sort that we now turn.

2. Restoring the commonsense concept of person and self to psychology as a science

Because the requirements of philosophical and scientific analysis are often reductionistic, the concepts of person and self have been analyzed into numerous parts and functions both within and between disciplines (Martin & Barresi, 2000, 2006). This fragmentation seems to be an essential outcome of progress in knowledge about human beings and all of nature. However, even if for most purposes science must break apart what common sense unifies, for other purposes maintaining unified concepts is essential for understanding phenomena. In common sense, or what from a scientific perspective is often referred to as “folk psychology,” human beings take themselves to be persons and selves, and these beliefs must be understood as phenomena, even if, in some philosophical or scientific ontologies, there are no real entities in the world that are persons or selves.

Whether persons and selves are unified objects or mere organizations of elemental material substances or psychological processes, in common sense human beings take themselves to be unified agents and this belief in personal unity is an apparent cause of much of their behavior. The fact that humans believe themselves to be persons and selves makes a difference to how their lives are lived. They take ownership of past acts that they remember doing, which leads them to act in ways consistent with past promises and intentions, and they often act with respect to goals to be achieved only in the distant future. Moreover, when they don't take ownership of their acts by keeping promises or acting consistently in terms of future goals, other humans treat them differently and typically in a manner that can have negative impact on their lives. While there are exceptions – for instance children and demented seniors – most humans can act in the present based on intentions sometimes made years ago for goals that can only be achieved in the distant future. For humans, this kind of long-term consistency in actions is not based primarily on habit, but on choice and intention. So, any theory of human nature and its activity must take into account these long-term, time-binding relations in human activity. The concept of a self-conscious person, who remembers his or her past, anticipates his or her future, and makes relevant time-binding decisions in the present, thus plays an essential role in accounting for what humans as biological and social organisms are about. Without this capacity for self-conscious personhood, we would not be the kind of organisms that we are. So even the needs of good science require us to take seriously the concepts of self and person as they apply to human psychology. Even if our self-conceptions are wrong, if they are persistent and ubiquitous, they are an important part of who we are, and hence part of what needs to be understood in order to understand ourselves.

3. An integrated theory of person and self

In tracing the history of theories of self and personhood briefly summarized above, we saw how analysis and fragmentation have taken apart these concepts. How are we to put them back together in order to understand the phenomena of our commonsense psychology of self and personhood? Barresi and Martin (in press) suggest that, as a counterbalance to fragmentation in theories of self, there is a current need to adopt an integrative stance toward person and self that takes into account several dimensions that have proved important in any discussion of these terms, which they call the *ontological dimension*, the *experiential dimension*, and the *social dimension* (see Martin & Barresi, 2006, especially chaps. 12–14; Martin, 2003; Seigel, 2005; Wiggins, 1987). Each of these dimensions focuses on an aspect of person and self that can be viewed as independent of the others, but can also be seen to be interdependent. Each typically plays a central and sometimes singular role in theories of self and person, often in different disciplines.

The ontological dimension focuses on what kind of thing a person or self is. On the commonsense view proposed here, self and person are two terms that we apply to the human organism as a whole through most if not all of its life. The focus of the experiential dimension is on the first-person experience of self, not with regard to what we can infer about what sort of thing, other than experience, the self might be, but on what our experience of the self is actually like. For whatever else we are, as humans we seem to experience ourselves in certain characteristic ways that often differ radically from the ways that non-human animals apparently experience themselves. An integrated account of person and self should take note of these characteristic ways in which humans experience themselves. The

third dimension is the social dimension. Individual humans relate socially to each other in unique ways unparalleled in any other species. Included in this dimension is the general role that social interaction beginning in infancy plays in self-awareness and in the developmental origins of self-concepts. But the social dimension also includes narrative accounts of self-identity that are the focus of personology, as well as social responsibilities, practical ethical principles, and ideals of the good life that we eventually form as adults. Ultimately, the social dimension is essential to our understanding of our selves as selves and as persons and is the basis for the rich understanding of self that is unique to our species.

In this theoretical framework, the human organism is acknowledged to be the primary source of unity for persons and selves. Under normal circumstances, the life history of the biological individual gives the person or self not only an objectively verifiable, unique identity that persists through time, but also a site for a variety of processes involving the other two dimensions of self to occur. This cannot be said for the other two dimensions. Under normal conditions, both the experiential and social dimensions of self depend on the existence of organic individuals as the basis for the self as defined by these other dimensions. In addition, the issue of self-unity is less problematic if the organic dimension is taken as primary than if the experiential or social dimensions are taken as primary. Because the person as self is taken as a unified organic individual extended in time, the main issue in relating the ontological dimension to the other dimensions is how that individual experiences its selfhood, in particular, how the first-person perspective of self develops in the individual's life history and the role that social existence and other individuals play in the development of self-consciousness.

There may be better ways to define the dimensions of self and person that need to be brought together to form an integrated theory. However, the present proposal shows that there is nothing inevitable about the theoretical fragmentation of the notions of self and person. Moreover, it indicates that an integrated framework of person and self can be developed, which is more in line with common sense and which can be used to understand human phenomena involving these concepts.

4. The phylogenesis and ontogenesis of persons and selves

Having described a general framework for investigating from a scientific point of view our commonsense experience of ourselves as persons and selves, we can now turn to possible applications of that framework. From a phylogenetic perspective, our primary interest in this section is to layout as clearly as possible what differentiates human experience of themselves as persons and selves from comparable experiences in other organisms. With respect to ontogenesis, the goal is to indicate when various aspects of our human sense of self and personhood develop during our lifetime.

In contrast to some theoretical interpretations, I have suggested that our commonsense notions of person and self are intimately linked so that each individual human is seen both as an embodied person and a self. Self and other are conceived equivalently as persons and selves that have both a first-person view of themselves and a third-person view of other individuals, but with a common conception of what these individuals are, which is a person. In considering personhood from a phylogenetic perspective, the key issue then is when does a common conception of self and other as individual embodied agents of the same kind with a first-person perspective of themselves first appear? And why? If we look at currently existing organisms, it turns out that there are very few that even come close to this possibility. Barresi and Moore (1996, in preparation) argue that most animals have quite distinct conceptions of self and other and that only very few have something akin to our notions of person and self. The problem is that any notion of self that most organisms have is conceived only from a first-person perspective. They do not experience or imagine themselves from a third-person perspective. They are always in the midst of their activity from a subjective viewpoint and cannot conceive of their activity from an objective viewpoint or the perspective of another individual. Furthermore, the converse is also true. They cannot imagine the first-person perspective of other individuals, but instead, always perceive others from a third-person perspective, as animate objects in their environment. The capacity to shift perspectives from that of one's own to that of another individual is not within the cognitive powers of most organisms. But there appear to be a few exceptions.

One form of evidence indicating the possibility that an animal can imagine itself from a third-person perspective, or has an objective concept of itself, appears when it is able to recognize itself in a mirror. Of particular importance as evidence of self-recognition is the capacity to recognize a mark

placed on a body part hidden from normal vision and to respond to it when placed before the mirror (Gallup, 1970). While this test does not guarantee a concept of itself equivalent to that it has of others, it does indicate, along with the use of mirrors to recognize other individuals, that the animal can treat the reflections in a mirror of itself and another individual equally as indicating a current state of a particular individual. Perhaps surprisingly, there are very few animals that can recognize their full bodies in a mirror, though many can recognize other individuals in the mirror, and use the mirror for other purposes involving particular body parts (e.g., Anderson & Gallup, 1999; Gallup, 1970; Mitchell, 2002).

Animals that pass stringent forms of the mirror test include great apes (gorillas, orangutans, chimps, and bonobos; Anderson & Gallup, 1999), cetaceans (e.g., dolphins and killer whales; Marino et al., 2007), elephants (Plotnik, de Waal, & Reiss, 2006), and at least one species of birds (magpie; Prior, Schwarz, & Gunturkun, 2006). All of these species have relatively large brains. Moreover, there is evidence in most of these species that they have the converse capacity to imagine the first-person perspective of others, something that again is quite unusual among most animals. For instance, great apes, cetaceans, and elephants show fairly strong evidence of empathy not only responding to the expressed distress of kin, but also to unexpressed situational needs of non-kin (Connor & Norris, 1982; Trivers, 1985; de Waal, 2008). There is also evidence to suggest that great apes and dolphins can imagine the visual viewpoint and appreciate knowledge of others (Call & Tomasello, 2008; Marino et al., 2007). Taken together, these findings suggest that such animals have comparable conceptions of self and other, conceiving of each from both a first- and a third-person point of view. Barresi and Moore (1996, in preparation) suggest that the conjunction of these two types of skills implies that such animals have a concept of an embodied agent that engages in intentional or psychological relations with other objects – a concept that includes both inner, or first-person, information about the agent's psychological orientations to objects as well as outer, or third-person, information about those orientations.

Although this concept of an individual embodied agent that these organisms seem to have may not be the same as the human concept of person, particularly as it occurs in later human development, it does come close to matching the concept of person that is available to a human 2-year-old. With this primitive concept, one is able to conceive of self and other as equivalent embodied agents that persist through time, each having its own unique and different perspective and intentional orientation on the same events or objects. However there are differences between humans and these other animals that already appear in comparison with 2-year-olds. While toddlers can engage in referential communication with gesture and language, without extensive training by humans these other animals do not. An important outcome of the use of referential language is the capacity to use the deictic pronouns “I” and “you” interchangeably between individuals, depending on perspective. The use of these terms provides strong evidence of a uniform conception of self and other as embodied intentional agents, particularly when combined with an appreciation of differences in point of view, for instance in preferences (e.g., Repacholi & Gopnik, 1997).

Two-year-olds also acquire knowledge of diverse cultural roles that are reversible between individuals, and they engage in more complex forms of imitation and play. Through extensive triadic interactions in which an adult and infant take turns manipulating objects and imitating each other's actions, 1-year-old infants become acquainted with reciprocal roles in shared activity. By the end of the second year, they can engage in simultaneous mutual imitation where two infants will copy each other's actions with objects, taking turns doing what the other does, and adopt diverse roles in their play activity. In contrast to 2-year-olds, attempts to show that chimps encode for later use complementary roles in joint tasks have generally failed to show transfer (Tomasello, Carpenter, Call, Behne, & Moll, 2005). So there are already differences early on between human concepts of person and the comparable notions that even large-brained animals have of self and other.

Moreover, there are other important differences that separate humans from other animals in their concepts of self and other. Among the things that both other animals and 2-year-olds lack is a concept of self and other that is a subject of private experience and thought, and an understanding that several individuals can have different phenomenal representations of a single event or object. They also lack a conception of a temporally extended self of the sort proposed by Locke when he defined personal identity in terms of consciousness, where one can recollect experiences and a point of view that one had in the past, anticipate experiences and a point of view that one will have in the future, and

differentiate them from one's present experience and differing point of view. These concepts come on line in humans around the age of 4 (Barresi, 2001), when they acquire what has been called a representational theory of mind (Perner, 1991) and an extended concept of self (Moore & Lemmon, 2001), and there is little evidence that any other animals possess similar concepts that can be applied equally to self and other (Suddendorf & Corballis, 2007). Although there are further differences between animals and humans in their conceptions of self and person, one of particular importance being the development in late adolescence of an abstract narrative conception of self, there is no need here to go into these additional differences.

5. Reciprocal altruism and the evolution of the human concept of person

How did our human conception of person evolve out of earlier primate concepts? I think that the simple answer is reciprocal altruism (Trivers, 1971). With the possible exception of magpies, all the species that have sufficient reflective capacity to recognize themselves in a mirror, show the beginnings of a form of altruism that occurs between individuals who are recognizable as individuals but who are not kin, and which takes into account the viewpoint of the other individual even in the absence of explicit signs of distress (de Waal, 2008). This implies an interest in the concerns of others who do not contribute directly to one's reproductive success, but who might do so indirectly through some form of reciprocal altruistic activity, either individually or through alliances. If there were not some long-term functional expectation of reciprocity, such activity, and the psychological mechanisms upon which it depends, would not occur. While it may be useful to understand the perspective of others and be able to self-reflect for competitive purposes, I believe that the coordination of these capacities in a concept that can be applied equally to self and other is more likely to evolve in circumstances requiring cooperative activity and empathy among non-kin. So, I believe that the conditions that gave a selective advantage to increasing levels of within group cooperation among non-kin and to more complex forms of reciprocal altruism within the hominid line, provided the evolutionary basis for our human concepts of person and self.

Essential for these human concepts of self and person is the capacity to represent the points of view of self and other equally. For the purposes of competition, having what is called a "Theory of Mind" (Premack & Woodruff, 1978) that is equally applicable to self and other is not required. It is certainly useful to anticipate the actions of others through a theory of their minds. And conversely, it might be useful to understand how others respond to one's own activity, using that theory. But this is not yet to conceive of self in the same manner as other. The theory is primarily about the other's mind and its activities, not about one's own. Moreover, even if there is some comparability in the theory of mind as it applies to self and other, in competitive circumstances the self is given priority over the other in all inferences that use the theory. So self and other as embodied agents engaged in psychological activity are not equated. The situation differs when cooperation is involved between genetically unrelated individuals. In cooperative relationships, where one must take into account another person's point of view and weigh it equally or almost equally to one's own, the ability to represent mental states of self and other equivalently becomes important.

According to Trivers' (1971) theory of reciprocal altruism, costs versus benefits for self and other must be calculated in a manner that will maintain equity across conditions and through time. If an act is of little cost to me and of great benefit to you, then I have a motive to perform it, so long as I can count on you to behave similarly in reversed circumstances. This requires a common metric to calculate costs and benefits across individuals; otherwise, one individual might take advantage of another. In Trivers' (1971) presentation of the theory, the ability of individuals to calculate costs and benefits across self and other in a common metric appears to be assumed rather than explained. This would not be a problem, so long as the medium of exchange is the same and can be calculated easily using an objective metric. This may be the case for some restricted forms of reciprocation having immediate functional significance, for instance, food exchange between vampire bats (Wilkinson, 1984). While a theory of mind would not be required in such cases, it does appear to be necessary for reciprocal altruism in humans. In our case, the use of a common metric requires a theory of mind in order to judge how costs and benefits appear to different individuals and to the same individual across time. The use of such a theory is essential in our attempts to maintain equity in relative costs and benefits to self and

others in reciprocal altruistic relationships, where the medium of exchange is highly variable and depends on diverse individual purposes. Although we certainly have a bias in our own favor in our calculations of costs and benefits to self and other, which Trivers associates with 'subtle cheating', such cheating can only work in a context where gross differences in costs and benefits are easily detected, but minor deviations are not. To do this, we need a theory of mind, and associated concept of person, that applies equally well to self and other both at a time and across time.

Trivers (1985, 2002, pp. 16–17) himself realized not long after publishing his theory that it could account for our sense of fairness and justice. He had already postulated in the theory that friendship, sympathy, gratitude, guilt, reparation, trust, moralistic aggression, and other social emotions may have evolved as psychological mechanisms that track cost/benefit ratios of altruistic acts, but he didn't mention our sense of justice or fairness (Trivers, 2006). However, it seems likely that human moral psychology emerged as a result of the evolution of reciprocal altruism in hominid species. This moral psychology is deeply dependent on a concept of person that we can apply equally to self and others, where the perspectives of different individuals, including motivations and actions, can be interpreted using a common metric.

If one compares adult humans to great apes in their capacity for theory of mind, one of the main differences is in their understanding of different representational perspectives both at a time and across time. Chimps live mainly in the present. They do not act with regard to a distant future, or with respect to a distant past. In contrast, humans have a notion of self that is extended in time and they have the capacity to engage in mental time travel (Moore & Lemmon, 2001; Suddendorf & Corballis, 2007). This ability to imagine past and future events from different points of view applies to both self and other, and can also be applied to different representational points of view in the present (Barresi, 2001; Barresi & Moore, in preparation). These capacities are linked together in humans. Our understanding of false belief in others and changing beliefs in self, requires the same ability for imagining changes in perspective through time (Barresi & Moore, in preparation). Moreover, this capacity is linked to our ability to act based on future motives that may be in conflict with current motives (Barresi, 2001; Moore, Barresi, & Thompson, 1998; Moore & Lemmon, 2001; Thompson, Barresi, & Moore, 1997). Whereas a chimp can act now for goals that will be achieved in the future, it requires a current motivation for the future goal. By contrast, a 4-year-old child can delay gratification of a current desire in favor of a desire it knows it will have in the future (Thompson et al., 1997). The child can also remember past episodes with different motives and knowledge as distinct from those he or she has in the present. Thus, the 4-year-old is capable of representing changes in representational perspective for self over time. And the child's capacity to do this is linked to a similar capacity to represent diversity in the perspectives of others. Thus, in congruence with the need of a common metric for representing diverse perspectives that is required for reciprocal altruism of a general sort to occur, humans have it at an early age, while chimps – and apparently other animals – never acquire it.

This is not to say that only humans exhibit reciprocal altruism. Limited forms of reciprocation may occur in diverse species (Trivers, 2006). With respect to more complex forms similar to humans, we have already noted chimpanzees, and other species that pass mirror tests, engage in some empathic acts of altruism with non-kin that appear to be a form of reciprocal altruism. Furthermore, in the wild, chimps and other primates show evidence of cooperation and reciprocal exchange of services of various kinds and even seem to act based on some sense of fairness or justice (de Waal, 2006). However, recent experimental studies attempting to test the theory of reciprocal altruism suggest that chimps have no regard for outcomes to others, but, instead, are only concerned with their own self-interest (Jensen, Hare, Call, & Tomasello, 2006). Furthermore, in studies where exchange of favors involves role reversal and turn taking across time, experimental evidence of contingent responding by primates to altruistic and non-altruistic partners proves to be quite limited (e.g., Brosnan et al., 2009; Melis, Hare, & Tomasello, 2008). Recently, Stevens and Hauser (2004) have argued that only very narrow forms of reciprocation occur in non-human cases because most animals lack the mental capacities necessary to deal with costs and benefits extended across time. But the evidence also suggests that they do not take into account the different perspectives of self and other, when conceiving of costs and benefits, and to compare them to each other (e.g., Jensen et al., 2006). Thus recent experimental findings are more congruent with the present proposal that what is unique about human reciprocal altruism is its dependence on a uniform concept of person that can be applied

equally to self and other, which involves a common metric for calculating costs and benefits across diverse perspectives, and which extends across time. This system makes it possible to calculate costs and benefits relative to different participants who must track each other's changing desires across time, while their commitments to each other are expected to persist over time. With this system in place, long-term relationships involving reciprocal altruism of various kinds become possible and a sense of justice becomes a medium by which to assess cost/benefit values of those relationships.

6. Moral psychology and our concept of person

William Hazlitt (1805) (Barresi, 2001; Martin & Barresi, 1995) wrote of imagination that: "it must carry me out of myself into the feeling of others by one and the same process by which I am thrown forward as it were into my own future being and interested in it. I could not love myself if I were not capable of loving others. Self-love, used in this sense, is in its fundamental principle the same with disinterested benevolence" (p. 3). In this proposal, Hazlitt suggested that both altruism with respect to others and prudence with respect to our selves have as a necessary requirement our ability to imagine motivations and perspectives that differ from that which we currently have. In the same book, he proposed that our notion of our selves as persons extended in time is an acquired concept, thereby providing one of the earliest developmental accounts of our sense of personal identity through time. With respect to moral psychology, it is interesting that he linked altruism to both prudence and personal identity and that both depended on a capacity for empathic imagination that can apply equally well to others as to self.

Much later, Nagel (1970) proposed that the "possibility of altruism" depends on a concept of person that applies equally to self and other in much the same manner as prudence depends on a concept of self that applies equally to a future and present self. According to Nagel, rational justification for prudent behavior relative to future motivations of self is linked to rational justification for altruistic behavior involving the motivation of others. In both cases a kind of neutral weighing of motivations in a common metric is involved. When considering altruism, Nagel uses the example of pain and role reversal to get his intuitions started. Whether you step on my foot or I step on your foot, the pain involved justifies not doing it regardless of whose pain it is. Hazlitt viewed pain in a similar allocentric manner. With respect to prudence, Nagel thought that one must believe in one's own personal identity through time in order to rationally justify actions now for motivations that one will have in the future. Thus, for both Hazlitt and Nagel, belief in personal identity and prudence for self as well as altruism for other persons are intimately linked.

John Rawls was Nagel's supervisor at Harvard, where Trivers' was also doing graduate work when he wrote his theory. In his *Theory of Justice* (1971), Rawls provides a different notion of moral psychology from either Hazlitt or Nagel, which focuses instead on justice as fairness. However, Rawls also cites Nagel's work with respect to prudence and criteria of personhood, as well as Trivers' theory, as providing indirect evolutionary support for his theory. Rawls proposed an idealized method for generating a just society by conceiving of individuals as having to make decisions about the structure of distribution of goods and services in a society "under the veil of ignorance" of who they would turn out to be in that society. I believe that Hazlitt, Nagel, and Rawls, as well as Kant in his categorical imperative, are all attempting to express notions of equality and equity that they derived from the sense of justice that is a concomitant of human understanding of self and other as persons and that has its origins in the evolution of reciprocal altruism in hominids. This notion of justice is also instantiated in the norms of reciprocity found throughout the world and in all major religions. It is fortunate that we have this concept, even if we often restrict its use to very narrow groups with which we identify (Barresi, 2004) and often view outsiders as non-persons that are less than human. Our notion of personhood is a fragile commodity – a recent evolutionary invention – something that we can, but do not automatically, apply equally to our selves and others.

7. Conclusion

John Locke's psychological theory of self and person, with its focus on consciousness and personal identity, captures an important part of our commonsense notion of person, and also an important

aspect of what makes humans different from other organisms. This aspect is the experience of our selves and other persons as persistent subjects of experiences extended in time. But, with respect to the integrative theory proposed here, what Locke left out was a direct link to embodiment and to social life. Darwin's theory provides an answer to the embodiment issue by giving a naturalistic account of humans as animals. What remains is to provide an account of social and moral personhood and the role of evolution in our human concept of person. I believe that reciprocal altruism can account for why hominids acquired a concept of self and person that represents diverse perspectives of self and others equally. It also provides a reason for our human belief in the persistence of self in time, including even into an afterlife. As Locke himself recognized, person is a term that is used when one is concerned with justice. But rather than mainly insuring that justice occurs in the afterlife as he thought, it also insures that long-term social relationships involving equity and justice occur in this life. Though only a recent invention and unique to humans, this concept of person and self is the source of our highest notions of human dignity and moral value. So, the least that we as psychologists can do is to give it some credence as an important concept for understanding what may be Mother Nature's finest gift to the human species and a unique aspect of human psychology.

Acknowledgements

This paper could not have been written without the prior and current collaborative work with Chris Moore and Raymond Martin. I also wish to thank SSHRC of Canada and Dalhousie University for research grant support.

References

- Anderson, J. R., & Gallup, G. G., Jr. (1999). Self-recognition in nonhuman primates: past and future challenges. In M. Haug, & R. E. Whatlen (Eds.), *Animal models of human emotion and cognition* (pp. 175–194). Washington, DC: American Psychological Association.
- Barresi, J. (1999). On becoming a person. *Philosophical Psychology*, 12, 79–98.
- Barresi, J. (2001). Extending self-consciousness into the future. In C. Moore, & K. Lemmon (Eds.), *The self in time: Developmental perspectives* (pp. 141–161). Hillsdale, NJ: Erlbaum.
- Barresi, J. (2004). Intentional relations and divergent perspectives in social understanding. In S. Gallagher, & S. Watson (Eds.), *Ipseity and alterity: Interdisciplinary approaches to intersubjectivity* (pp. 74–99). Rouen: Presses Universitaires de Rouen.
- Barresi, J. (2006). On earth as it is in heaven: Trinitarian influences on Locke's account of personal identity. *The Pluralist*, 1, 110–128.
- Barresi, J., & Juckes, T. (1997). Personology and the narrative interpretation of lives. *Journal of Personality*, 65, 693–719.
- Barresi, J., & Martin, R. History as prologue: western theories of the self. In S. Gallagher (Ed.), *Oxford handbook of the self*, in press.
- Barresi, J., & Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences*, 19, 107–122.
- Barresi, J., & Moore, C. *Understanding intentional relations of self and other*. Oxford, England: Oxford University Press, in preparation.
- Brosnan, S. F., Silk, J. B., Henrich, J., Mareno, M. C., Lambeth, S. P., & Schapiro, S. J. (2009). Chimpanzees (*Pan troglodytes*) do not develop contingent reciprocity in an experimental task. *Animal Cognition*, 12, 587–597.
- Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Science*, 12, 187–192.
- Connor, R. C., & Norris, K. S. (1982). Are dolphins and whales reciprocal altruists? *American Naturalist*, 119, 358–374.
- Erikson, E. H. (1980). *Identity and the life cycle*. New York: Norton (Original work published in 1959).
- Gallup, G. G., Jr. (1970). Chimpanzees: self recognition. *Science*, 167, 86–87.
- Hazlitt, W. (1805/1969). *An essay on the principles of human action and some remarks on the systems of Hartley and Helvetius*. Reprint, intro. J. R. Nabholz. Gainesville, Fla.: Scholars' Facsimiles & Reprints.
- Jensen, K., Hare, B., Call, J., & Tomasello, M. (2006). What's in it for me? Self-regard precludes altruism and spite in chimpanzees. *Proceedings of the Royal Society B: Biological Science*, 273, 1013–1021.
- Locke, J. (1694/1975). *An essay concerning human understanding*. Oxford: Clarendon Press.
- Marino, L., Connor, R. C., Fordyce, R. E., Herman, L. M., Hof, P. R., Lefebvre, L., et al. (2007). Cetaceans have complex brains for complex cognition. *PLoS Biology*, 5, 0966–0972.
- Martin, J. (2003). Emergent persons. *New Ideas in Psychology*, 21, 85–99.
- Martin, R., & Barresi, J. (1995). Hazlitt on the future of 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.
- Martin, R., & Barresi, J. (2006). *The rise and fall of soul and self: An intellectual history of personal identity*. New York: Columbia University Press.
- McAdams, D. P. (1990). Unity and purpose in human lives: the emergence of identity as a life story. In A. I. Rabin, R. A. Zucker, R. A. Emmons, & S. Frank (Eds.), *Studying persons and lives* (pp. 148–200). New York: Springer.
- Melis, A., Hare, B., & Tomasello, M. (2008). Do chimpanzees reciprocate received favours? *Animal Behaviour*, 76, 951–962.
- Mitchell, R. W. (2002). Subjectivity and self-recognition in animals. In M. R. Leary, & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 567–593). NY: Guilford Press.

- Moore, C., Barresi, J., & Thompson, C. (1998). The cognitive basis of future-oriented prosocial behavior. *Social Development*, 7, 198–218.
- Moore, C., & Lemmon, K. (Eds.). (2001). *The self in time: Developmental perspectives*. Hillsdale, NJ: Erlbaum.
- Murray, H. A. (1938). *Explorations in personality: A clinical and experimental study of fifty men of college age*. New York: Oxford University Press.
- Nagel, T. (1970). *The possibility of altruism*. Oxford: Oxford University Press.
- Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.
- Plotnik, J. M., de Waal, F. B. M., & Reiss, D. (2006). Self-recognition in an Asian elephant. *Proceedings of the National Academy of Sciences*, 103, 17054–17057.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1, 515–525.
- Prior, H., Schwarz, A., & Gunturkun, O. (2006). Mirror-induced behavior in the magpie (*Pica pica*): evidence of self-recognition. *PLoS Biology*, 6, 1642–1650.
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.
- Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: evidence from 14- and 18-month-olds. *Developmental Psychology*, 33, 12–21.
- Rogers, C. (1955). Persons or science: a philosophical question. *American Psychologist*, 10, 267–278.
- Rogers, C. (1961). *On becoming a person*. Boston: Houghton Mifflin Company.
- Seigel, J. (2005). *The idea of the self: Thought and experience in Western Europe since the seventeenth century*. Cambridge: Cambridge University Press.
- Stevens, J. R., & Hauser, M. D. (2004). Why be nice? Psychological constraints on the evolution of cooperation. *Trends in Cognitive Sciences*, 8, 60–65.
- Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: what is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, 30, 299–351.
- Thompson, C., Barresi, J., & Moore, C. (1997). The development of future-oriented prudence and altruism in preschool children. *Cognitive Development*, 12, 199–212.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675–735.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46, 35–57.
- Trivers, R. L. (1985). *Social evolution*. Menlo Park: Benjamin/Cummings Publishing Company, Inc.
- Trivers, R. L. (2002). *Natural selection and social theory: Selected papers of Robert L. Trivers*. In *Evolution and cognition series*. Oxford: Oxford University Press.
- Trivers, R. L. (2006). Reciprocal altruism 30 years later. In P. M. Kappeler, & C. P. van Schaik (Eds.), *Cooperation in primates and humans: Mechanisms and evolution* (pp. 67–83). Berlin: Springer-Verlag.
- de Waal, F. B. M. (2006). Joint ventures require joint payoffs: fairness among primates. *Social Research*, 73, 349–364.
- de Waal, F. B. M. (2008). Putting the altruism back into altruism: the evolution of empathy. *Annual Review of Psychology*, 59, 279–300.
- Wiggins, D. (1987). The person as object of science, as subject of experience, and as locus of value. In A. Peacocke, & G. Gillett (Eds.), *Persons and personality: A contemporary inquiry* (pp. 56–74). Oxford: Basil Blackwell, Ltd.
- Wilkinson, G. S. (1984). Reciprocal food sharing in the vampire bat. *Nature*, 308, 181–184.