Children’s Theories of the Self

Christina Starmans
University of Toronto

This article provides a theoretical review of the developmental origins of children’s “folk theories” about the nature of the self, linking theoretical developments in philosophy with empirical discoveries from developmental psychology. The article first reviews children’s views about the material nature of the self, outlining evidence that children naturally think about the self as distinct from the body. It then discusses children’s understanding of the persistence of the self over time and, finally, explores children’s views about conflict within the self. Together, these findings suggest that preschoolers possess stable, coherent, and predictive theories about the nature of the self that are stable across individuals, early emerging, and in some cases undergo interesting developmental change.

Children, like adults, spend a great deal of time thinking about themselves. Even infants seem to have at least a rudimentary sense of self, and as children get older, they develop more sophisticated mental representations of themselves as individuals, attributing distinctive physical and psychological traits to themselves and others, evaluating themselves through social comparison, and using their self-views to navigate social relationships. A substantial body of research suggests that these developing self-views have a substantial impact on children’s social relationships, mental health, and academic performance (e.g., Bandura, 1997; Dweck, 2006; Kochanska, Koenig, Barry, Kim, & Yoon, 2010).

Most discussions of how children begin to represent and reason about the self involve thinking about the unique characteristics that makeup a self—those characteristics that make me me and you you. However, these discussions tacitly build upon an unexamined assumption: that children have a concept of what it means to have a self at all. If the important topics described earlier are all circling around the broad question “Who am I?,” the question here is instead “What am I?” Thinking about the self in this way involves considering what all selves have in common—that is, what is a self, who has one, and why does it matter?

These are some of the most difficult questions about the nature of being human, and they have often been seen as the purview of philosophers. However, underlying complex arguments about the metaphysical nature of the self is an equally interesting psychological question. How do people naturally think about what a self is, and what are the developmental roots of this commonsense conception?

It is worth considering that there may be no answer to the question of how children naturally think about the self in this sense. Thinking about the nature of the self might be like thinking about quantum physics: only experts really have any idea how to think about it, and most people, especially children, may never have considered it. However, people who have never reflected on the nature of the self possess at least a minimal theory about the self when they, for example, assume that a person who has been in a disfiguring accident is still the same self even after losing large parts of her body (suggesting they do not see a person as equivalent to her body), or assign a person moral responsibility for actions that he took long in the past (suggesting they view the self as persisting across time and change). This suggests that there may be some basic intuitions about the nature of the self that...
underlie fundamental aspects of social cognition, and thus we might expect these intuitions to be present very early in development.

A classic focus of developmental psychology has been to examine the ways in which children develop “folk theories” to help them make sense of sophisticated concepts like this (Carey, 1988; Keil, 2010). For example, even young infants have an intuitive sense of physics, expecting objects to be solid, to fall down if not supported, and to exist continuously through space and time (Baillargeon, 2004; Spelke, 1994). Likewise, 8-month-olds demonstrate a “folk biology,” expecting animals to have insides instead of being hollow (Setoh, Wu, Baillargeon, & Gelman, 2013), and 3-year-olds know that animals, but not artifacts, can grow, heal themselves, and reproduce (Backscheider, Shatz, & Gelman, 1993; Hatano & Inagaki, 2013). Perhaps most relevant, a large body of work has investigated children’s “folk psychology”—how children understand the beliefs and desires of others (e.g., Wellman, 1990). These intuitive theories have been characterized as domain-specific sets of beliefs that arise informally and without formal instruction, and which are, at least to some degree, coherent, consistent, and predictive (Carey, 1985; Carey & Spelke, 1996; Keil, 2010; Slaughter & Gopnik, 1996). In this article, I suggest that children may likewise have intuitive theories about the nature of the self.

Understanding the origins of intuitive theories of the self is essential. For one thing, these intuitions are the seeds that later develop into complex philosophical and scientific lines of inquiry, and form the basis of many religious beliefs. People’s conception of the self impacts their views about numerous social and political issues such as organ donation, the insanity defense, and abortion debates. Attributing a self—or certain features of a self—to an entity affects the moral rights and responsibilities we confer upon that entity (e.g., Gray, Gray, & Wegner, 2007), and thus reasoning about the moral significance of fetuses, chimpanzees, sophisticated robots, or people in a vegetative state rests in part on beliefs about the nature of the self.

Research into the developmental origins of intuitive theories about the self is still in its infancy. The goal of this article is to review the mounting evidence suggesting that children do indeed have stable, coherent, and predictive theories about the nature of the self, and to outline some key areas and open questions for future research. The article is organized around three big questions borrowed from philosophers of mind. The first concerns the material nature of the self: Do children, like adults, naturally see the self as distinct from the body? The second concerns the persistence of the self over time. How do children think about past and future versions of a self, and how can the self change over time while still remaining the same self? Third, do children see the self as a unified whole, or do they understand that a person may have multiple parts of the self that may be in conflict with each other? Across each of these questions, evidence is found for stable, early-emerging intuitions about the nature of the self, as well as some interesting cultural and developmental variation.

What is the Self?

Humans have long wondered about the nature of our selves. Are people simply physical objects, no different from rocks and trees? Or are we something more than that—are we immaterial souls, essentially different from these material things? Most adults around the world explicitly believe that they are something more than automata, propelled by physical stimuli to engage with other physical stimuli. Instead, they embrace a dualist view of the self, in which the self is separate from the body and might even survive its destruction—by living on in heaven or hell, or even occupying a different body, as in reincarnation. While the details differ, some have argued that the widespread belief in an immaterial self that is distinct from the body suggests that we naturally see the world in a way that causes us to categorize bodies and minds as two separate sorts of things (Bering, 2006; Bloom, 2004; Starmans & Bloom, 2011). If so, we should expect to see evidence of a dualist worldview even in the judgments of young children.

Research in developmental psychology suggests that this may indeed be the case. From a very young age, infants make important distinctions between objects with selves and objects without selves. Cues such as the presence of faces and eyes, asymmetry along one axis, nonrigid transformation such as expansion or contraction, self-propulsion, and the ability to engage in contingent and reciprocal interactions with other agents lead infants to classify something as a psychological agent (Johnson, 2000). Infants then make different predictions about these agents than they do about objects. For example, infants expect agents, but not objects, to have consistent goals and preferences (Woodward, 1998), to act rationally (Gergely & Csibra, 2003), and to create order out of disorder (Newman, Keil, Kuhlmeier, & Wynn, 2010). Infants also make different judgments
about behavior toward agents than behavior toward objects. For example, 6-month-olds like an agent who helps another agent up a hill, but are neutral toward an agent who pushes an object in the same manner (Hamlin, Wynn, & Bloom, 2007). This fundamental distinction between agents and objects may underpin dualist intuitions by forcing infants to understand the world through two distinct cognitive systems (Bloom, 2004).

Once children are old enough to verbally report their beliefs, researchers can investigate their early intuitions about selves. Given that many adult beliefs about the conscious self relate to the continuation of life after death, one place to start asking these questions is to ask what children think happens when someone dies. In one such study, Bering and Bjorklund (2004) told 4- to 6-year-old children a story about a mouse that gets eaten by an alligator, and then asking what would happen to the mouse now that he was no longer alive. Children knew that the mouse’s biological processes like growing older or needing to eat would cease. However, they thought that psychological processes such as beliefs, desires, and emotions would continue after death. This suggests that preschoolers intuitively separate the mind from the body, and assume that some part of a living being can survive the death of the body. These early intuitions appear to be independent of parents’ views of death, one biological in which most or all mental processes cease and one religious in which mental processes continue after death. This pattern of results as suggesting that children between the ages of 4 and 6 have a dualist view of the self, characterized by the belief that the self is a substance whose entire essence, or nature, consists in thinking. For Descartes, this immaterial nature of the self suggests that it “need have no location.” However, recent research suggests that this is an interesting way in which commonsense dualism departs from Descartes’ dualism.

Where is the Self?

The claim so far is that children think that they are immaterial; they are not identical to their physical bodies. Like Descartes (1637/1965, p. 28), they feel as though they are “a substance whose entire essence, or nature, consists in thinking.” For Descartes, this immaterial nature of the self suggests that it “need have no location.” However, recent research suggests that this is an interesting way in which commonsense dualism departs from Descartes’ dualism.

When adults are asked explicitly about the location of their self, most indicate that it is located within the head (Bertossa, Besa, Ferrari, & Ferri, 2008; Limanowski & Hecht, 2011) or in the chest (Limanowski & Hecht, 2011). However, these judgments may be due to any number of factors, including explicitly held religious, cultural, or scientific beliefs. Aside from these cultural beliefs, is there an intuitive sense of the self as being located in a particular spot in the body? As reviewed earlier, young children have had less opportunity to acquire religious and scientific ideas about the self and have a limited understanding of the brain as the producer of mental life (Flavell et al., 1995; Gottfried et al., 1999; see Bloom, 2004 for review). Where do they think that the self is?

There is some indirect evidence to suggest that, in fact, preschool-aged children do have intuitions about the location of the self. In a study investigating children’s egocentrism, Flavell, Shipstead, and Croft (1980) found that when 2.5- to 4-year-old children had their eyes covered, they judged that an experimenter could not see them. However, they acknowledged that the experimenter could see their arm, despite the fact that the child themselves could not, suggesting that they were capable of taking the experimenter’s perspective at least some of the time. Likewise, 2- and 3-year-old children claimed that they could see a doll if its legs were occluded, but that they could not see her if her head was occluded (McGuigan & Doherty, 2006; Russell, Gee, & Bullard, 2012), and 3-year-olds deny that they can “see” a person whose eyes are covered, but affirm that they can “look at” the person (Moll, Arellano, Guzman, Cordova, & Madrigal, 2015).

The researchers interpreted this curious pattern of results as suggesting that children between the
ages of about 2 and 4 may have a different notion than adults of what it means to “see” a person, suggesting that children may misconstrue the concept of “seeing” a human target as an activity that requires mutual engagement (McGuigan & Doherty, 2006; Moll et al., 2015; Russell et al., 2012). However, an alternate explanation for these findings may be that children intuitively see the head or the eyes as the “location” of the individual, and intend the claim that the experimenter “can’t see me” to mean “the experimenter can’t see my self.”

One series of studies directly investigated the question of whether 4- and 5-year-old children naturally think of the self as being located in a particular part of the body, using an implicit method that asked children when an object was closest to a person (Starmans & Bloom, 2012). The logic of this method was that if children consider the self to be equally distributed across the body, or if they think the self has no spatial location, then they should judge that objects are equally close to a person regardless of where on her body they are positioned. However, if children have an intuition that the self is located in a particular part of the body, then they should judge that objects nearer to that part of the body are closest to the person. These studies found that 4- and 5-year-olds had a strong bias to report that objects were closest to a person when they were near her eyes (as opposed to her mouth, chest, stomach, or feet). This also held for an alien character whose eyes were located on his chest, and for objects located near a person’s eyes but out of her sight, suggesting that children were not simply considering which objects the person could see most closely. This suggests that children have a strong sense that the self is located near the eyes, even when the eyes are not located in the head. Importantly, children were not asked directly about the self or its location in these studies. The indirect nature of the methods used in these experiments suggest that children’s judgments do not result from a culturally learned understanding of the role of the brain in producing mental states, but might instead be rooted in a more intuitive or phenomenological sense of where in the body the self resides.

Research on adults’ implicit judgments about the location of the self has produced mixed results. Studies using the “closer object” method described earlier have found that adults share children’s intuitions that the self is located in the head, near the eyes (Starmans & Bloom, 2012). However, studies using other methods suggest that adults’ intuitions might be more variable. For example, when adults were asked to rotate a rod until it is pointed “directly at you,” about half the subjects pointed the rod toward their torsos, while the other half pointed it toward their heads (Alsmith & Longo, 2014). Which half a given participant fell into was largely determined by the initial position of the rod: If the rod arrived at the torso before the head, participants were much more likely to stop at the torso, and vice versa. This suggests that there may be as yet undiscovered factors that shift a person’s sense of the location of her self, and leaves open the question of whether children’s intuitions also vary in similar ways.

When is the Self?

The commonsense conception of a person as both a physical body and a psychological self not only affects where a person is thought to be, but also how that person is thought to persist over time. Throughout a person’s lifetime, both the body and the self change tremendously. Each person begins life as a miniature human, with much more in common with any other unrelated infant than with the wrinkled old person the infant eventually becomes. The baby and the senior citizen have none of the same memories, preferences, or knowledge, they look entirely different from one another, and they are not even made up of the same physical stuff, since most of the baby’s cells have died and been replaced by new cells. But if there is very little that is the same about the person at the end of one’s life and the person at its beginning, how do we identify them as the same person?

Older work suggests that children have a difficult time tracking identity across time. Carey (1985) suggested that children have a limited understanding of identity across growth and change until age 9 or 10. Six-year-olds place high importance on outward appearance and name, claiming that they would no longer be “the same person” if their name were to change, and that they were not now the same person they had been when they were a baby, and would no longer be the same person they are now once they grew older (Guardo & Bohan, 1971). When asked what aspects of themselves would need to change in order for them to become their best friend, or why characters in movies remained the same person throughout the movie, 6-year-olds again focused on superficial characteristics like physical appearance and name (Chandler, Lalonde, Sokol, Hallett, & Marcia, 2003; Mohr, 1978).

However, these findings may underestimate children’s understanding of the continuation of identity...
across time. Even for adults, the phrase “same person” is ambiguous—in some important ways, adults are not now the “same” person that they were as children, and a person whose name or appearance has changed is also reasonably thought of as being “different” than they were before. To avoid this ambiguity, Gutheil and Rosengren (1996) asked 4- and 5-year-olds whether the idiosyncratic characteristics of a particular animal (e.g., he will only eat out of the red bowl) would persist despite a change in name or appearance. By age 4, most children judged that these properties (and thus, presumably, identity) had persisted through the change. Similarly, 4-year-olds correctly judged that two identical Winnie-the-Pooh dolls who had access to different information would not share the same knowledge (Gutheil, Gelman, Klein, Michos, & Kelaita, 2008). These researchers and others argue that these findings provide evidence that children are using an individual’s spatiotemporal history—the physical path he or she traces through the world—to track identity (Gelman, 2003; Gutheil & Rosengren, 1996; Gutheil et al., 2008).

All of the strategies described earlier for tracking the identity of an individual person across time—name, appearance, spatiotemporal history—are the very same strategies adults and children use to track the identity of an individual object across time (Bloom & Markson, 1998; Gelman & Bloom, 2000; Xu & Carey, 1996). However, as discussed earlier, even young children and infants recognize that a key feature of humans and other agents is their internal, invisible, mental life—their selves. Many philosophers argue that someone remains the same person across her life span because there is a continuation of her conscious mind (e.g., Locke, 1690/1964). On this view, as on Descartes’, the essence of a person is psychological, and thus what makes a person the same person over time is the continuity of their mental lives. This view seems to capture some of our commonsense ideas about the self; adults agree that in the unlikely event that a person’s body and mind were to be separated, the person will be located wherever his or her mind is now located (Blok, Newman, & Rips, 2005; Corriveau, Pasquini, & Harris, 2005).

There is some evidence that young children also consider mental characteristics to be more associated with identity than physical characteristics. One study asked 5- to 6-year-old children to consider a “duplicating machine.” After seeing the machine copy some toys, children then saw a live hamster “duplicated” by the machine, and were asked about the physical and mental characteristics of the duplicate hamster. Children thought that the second hamster would share the unique physical properties of the original (having a marble in its stomach, a blue heart, and a broken tooth), but not the knowledge states of the original—it would not know that the child had recently told the hamster his or her name, tickled the hamster, and showed it a picture (Hood, Gjersoe, & Bloom, 2012).

This focus on knowledge (or episodic memory) as indicative of identity is consistent both with other work investigating identity judgments in children (e.g., Gutheil et al., 2008), and with a longstanding focus by philosophers on the continuation of episodic memory as central to personal identity. However, recent studies on adults’ intuitions about what makes someone the same person across time suggests that other psychological characteristics, such as personality, desires, preferences, and especially moral traits may also play a role (Strohminger & Nichols, 2014). To our knowledge, the developmental origins of these views have not yet been investigated, and thus it will be important for future studies to examine whether children also view changes in moral values, or other psychological changes, as disruptive of personal identity.

**Future Selves as Others?**

Both adults and children naturally think of a self as tracing a continuous path from a person’s birth to her death, or perhaps even persisting after death. Yet a growing body of evidence from social psychology, neuroscience, and behavioral economics suggests that there may be some interesting ways in which adults sometimes reason about temporally distant versions of themselves as if they were other people entirely—making more rational decisions, showing less generosity, and eliminating many well-known self-serving biases. For example, many well-documented self-other asymmetries—such as the “fundamental attribution error,” and adults’ tendency to see stable traits in others but situational variability in themselves—also turn out to be “current self–future self” asymmetries (see Fronin, 2008 for review). This tendency to view one’s future self as a third party is also reflected in differential brain activity. Adults recruit brain regions associated with “self” when thinking about their present selves, but brain regions associated with “other” when thinking about their temporally distant future selves (Mitchell, Schirmer, Amés, & Gilbert, 2011).

Less is known about the developmental origins of these self–other asymmetries in judgments of present and future selves. The logic of many of these studies with adults is twofold: first, demonstrate differences
or “biases” in how adults reason about others versus themselves, and second, examine whether these same biases appear when reasoning about present selves versus future selves. Neither step has yet been well documented in children. For example, no studies to our knowledge have investigated whether children show self–other differences in attributing behavior to internal enduring characteristics or temporary situational factors. This is partly because many early studies found that children do not begin to use character traits to explain others’ behavior until middle childhood (e.g., Peers & Secord, 1973). However, more recent studies have found that 3-year-olds use trait labels rather than physical appearance to predict the behavior of a story character (Heyman & Gelman, 1999), and 4-year-olds infer traits (e.g., nice, mean) from descriptions of behavior (Boseovski & Lee, 2006; Liu, Gelman, & Wellman, 2007). Most recently, Chen, Corriveau, and Harris (2016) found that preschoolers in both the United States and Taiwan spontaneously inferred character traits from behavior, and used these traits to make predictions about future behavior. Thus, it seems that children, like adults, do often explain the behavior of others in terms of internal enduring characteristics. However, no research to date has compared children’s propensity to explain the actions of others in terms of traits to their explanations for their own actions. Further investigations in this domain will be useful not only in service of documenting self–other differences in young children, but is also a crucial first step in an exploration of whether children might show self–other bias when considering their present selves and their future selves.

Other studies have taken a different approach to investigating how children and adults think about the self persisting across time. These studies asked 4- and 5-year-old children and adults about the ideal distribution of happiness across the life span of an individual—for example, is it better to have an equal amount of happiness every year, or a bit more happiness in some years, and less happiness in other years? A separate group of participants gave similar judgments for how happiness should ideally be distributed across a group of separate individuals. These studies found striking parallels in both preschoolers and adults’ judgments about the ideal life and the ideal society (Starmans & Bloom, 2013).

For example, children preferred lives with a mix of mildly happy years and mildly sad years over lives that had a uniform level of happiness. However, this preference depended on how extreme the variation was. When the varied lives included years of extreme happiness and extreme sadness, children were more likely to prefer a life of uniform moderately happy years. Intriguingly, children’s judgments about how best to distribute happiness across a group of separate individuals were affected by the very same factors: they preferred a population with a mix of mildly happy and mildly sad people over one where everyone is the same—but did not want a population with extremely happy and extremely sad people. These parallels also appeared for choices about sacrificing 1 year (or person) so that all the others would be happier, and for choices between short very happy lives (or small very happy populations) and long moderately happy lives (or large moderately happy populations).

While there are several possible interpretations of these parallel judgments, these data are at least consistent with the body of literature finding that adults sometimes think of their future selves as other people, and suggest that, in at least some circumstances, thinking about the life span of one individual as being composed of a series of separate selves (baby, toddler, teenager, adult, etc.) may be a natural intuition that is present very early in life.

Also consistent with this proposal are a number of studies finding that decision making for future selves more closely resembles decision making for others than for the present self. For instance, adults often prefer immediate gratification, even at the cost of long-term benefits. When making this same choice for someone else, however, adults choose the larger, but later, reward. Similarly, when choosing for one’s future self—for instance, when offered $50 two months from today, or $65 four months from today—adults show the same rational maximizing decision they made for others: they choose the $65 (Frederick, Loewenstein, & O’Donoghue, 2002; Pronin, Olivola, & Kennedy, 2008). Many theorists have argued that this is because adults are thinking of their future selves—the person who will be receiving the money in several months—as someone other than themselves (e.g., Ainslie, 1975; Metcalfe & Mischel, 1999; Thaler & Shefrin, 1981).

Several studies have found that children are even more likely than adults to show this “temporal discounting” (e.g., Green, Fry, & Myerson, 1994; Steinberg et al., 2009). This correlation between impatience and age likely has multiple causes, including an increasing ability to inhibit current desires (Metcalfe & Mischel, 1999; Mischel, Shoda, & Rodriguez, 1989), increasing experience with long delays (Green et al., 1994), an increasing ability to delay gratification (Mischel et al., 1989), and even a rational expectation that adults have a shorter future ahead of them than do children, and thus should
discount the future more steeply (Green et al., 1994). But as in the adult studies discussed earlier, a further possibility is that children see their future selves as other people, and prefer that their present selves have rewards, rather than these rewards going to someone else. If so, this would predict that children’s decisions about larger or smaller rewards for future versions of themselves (i.e., when all reward points are moved into the future) would, like adults, more closely resemble their decisions for others than decisions for their present selves. There is some evidence that children as young as 3 years of age make more patient decisions for others than for themselves (Prencipe & Zelazo, 2005), but little is known about children’s delay decisions for future versions of themselves. Thus, an important avenue for future research will be to investigate the relationship between children’s discounting decisions for others and for future versions of themselves.

For adults, the tendency to treat future selves as others is intensified when our future selves will have experienced identity-altering transformations, such as marriage, divorce, religious conversion, or even graduation from college (Bartels & Rips, 2010). One of the largest such transformations is, of course, development itself. This may suggest that children will be even more likely than adults to think about future selves as if they are other people, and may see even their near-future selves as less a part of their selves than adults. Consistent with this prediction, recent work investigating how similar 10- to 73-year-olds feel they are to their past selves, or will be to their future selves, suggests that the “temporal size” of the self increases with age (Starmans, Rand, & Bloom, 2016). That is, 10-year-olds feel that their selves extend only a short distance into the past and future, while older adults have a much larger window of time that they consider to be “me.”

A converging body of evidence thus suggests that both adults and children may, in some interesting ways, think about future versions of themselves as if they are other people. While research on adults’ reasoning about future selves is beginning to fill in, there remain many open questions as to when children may think of their future selves as they think of others, and what effects this may have on their decision making.

**Conflict in the Self**

The work discussed earlier aims to explore the idea that we think of our future selves as a series of different individuals. A related idea which has been discussed by many philosophers and psychologists is that even at a given moment in time, a person may feel as though she is composed of multiple selves. One of the earliest descriptions of this predicament is Plato’s idea that the soul is composed of three distinct parts: the appetitive part, consisting of base desires for alluring things like sweet and fatty foods or sex; the rational part, pulling us toward things that are good for us in the long run; and spirited part, ensuring that the dictates of reason are translated into action (Plato, 360 BCE/1955). Others have described these conflicting selves as a “hot,” emotional “System 1” versus a “cool,” rational, System 2 (Kahneman, 2011), the “ego,” the “superego,” and the “id” (Freud, 1923/1960), conscious versus unconscious processes (Barth & Morsella, 2008), or controlled versus automatic processes (Haidt, 2006). How do children reason about this inner conflict between different parts of the self?

Research examining children’s ability to identify conflicting mental states has produced mixed results. While some researchers find that children fail to understand that one person might have conflicting or mixed desires until the age of 7 or 8 (Hartner & Buddin, 1987; Lagattuta, 2005), others have found that with simpler tasks and more explicit forced choices 4- and 5-year-olds can identify mixed emotions (Kestenbaum & Gelman, 1995). Children begin to understand their own conflicting desires and develop strategies for self-control by around age 4 (Mischel & Ebbesen, 1970; Mischel & Mischel, 1983). And recent research has suggested that in a simple behavioral task, 3-year-olds understand that a person may have multiple desires that conflict with one another. For example, when an adult had a specifically stated desire to reach a specific cup, but an overarching desire to pour a glass of water to drink, 3-year-olds opted not to give her the requested cup if it was broken. Instead, children were able to override the request directed toward short-term goal, and provided a different, functional, cup that would meet the requester’s overarching goal. Not only did children hand over a different object than the one requested by the experimenter, but they often took pains to explain themselves to the adult, providing strong indication that they understood both desires clearly (Martin & Olson, 2013).

One common place that people experience inner conflict is in the domain of morality. This has long been recognized by moral philosophers. Aristotle, for example, argued that an act should only be considered moral if it is easy for the actor to do,
suggesting that a person’s inner harmony or conflict regarding their moral decisions should be central to moral judgments of that person. Conversely, Kant (1785) argued that an action is only truly moral if it is difficult—if the actor experiences some degree of inner conflict, yet chooses to do the moral action despite being tempted toward the immoral. Which, if either, of these moral philosophies capture children’s reasoning about inner moral conflict?  

Some findings suggest that both children and adults sometimes take an Aristotelian stance, emphasizing consistency in the self, or a person’s broader character, when making moral judgments (for adult studies, see Fleming & Darley, 1989; Pizarro & Tannenbaum, 2011). For example, preschool-aged children judge the same action to be more moral when the actor is described as globally nice, and the same bad action to be more immoral when the actor is described as globally mean (Jones, Tobias, Pauley, Thomson, & Johnson, 2009). And both adults and children have better memory for those who helped or hurt them when the action was out of character (i.e., if a typically mean person helped or a typically nice person harmed; Chang & Sanfey, 2009).  

However, other findings suggest that at least adults may sometimes take a more Kantian approach. For example, people who resist immoral desires are seen as more trustworthy, and the stronger the desire, the greater the credit for resisting it (Kee, 1969). Furthermore, adults sometimes prefer moral actors who exhibit conflicting mental states, such as reluctance to do a distasteful act while doing it for the greater good (Pizarro & Tannenbaum, 2011).  

Recent studies have directly investigated the role of inner conflict in children’s and adults’ moral judgments. In one study, participants were asked which of two characters was “more good,” when both characters performed the same moral action, but one experienced inner conflict and the other did not. While adults judged that the character who had to overcome the conflicting desire to do a bad action deserved more moral credit, 3- to 8-year-old children overwhelmingly chose the character with no inner conflict as “more good” (Starmans & Bloom, 2012). This suggests that children may begin with a more Aristotelian view of morality, approving most of those who do not struggle with moral decisions, but sometime between 8 years of age and adulthood develop a more Kantian framework, judging the worth of moral actions based on how difficult the action was for the actor.  

Might children reason differently than adults about inner conflict because they are less likely to experience it themselves? Although children certainly behave immorally, it is possible that they less often experience concurrent conflicting desires to do both the right and the wrong thing. However, navigating inner conflict, particularly in the moral domain, is part of everyday life from early childhood. The experience of wanting something but knowing that it is wrong (or at least not wanting to experience the consequences of the action) begins very early in life, with children beginning to be constrained by safety rules, property issues, and household rules in the early toddler years (Dunn & Munn, 1987; Kopp, 1982; Ross, Friedman, & Field, 2015). Preschoolers will consciously engage in tactics so as to distract themselves from immediate temptation in service of a competing goal (Mischel & Ebbesen, 1970), and will help others to delay gratification when they express conflicting desires (Prencipe & Zelazo, 2005). The ability of a preschooler to resist temptation is linked to greater social responsibility at the time, less aggressive behavior in middle childhood (ages 6–12; Rodriguez, Mischel, & Shoda, 1989), and a lower likelihood of drug use later in life (Ayduk et al., 2000). Thus, the developing understanding of conflict in the self is interesting not only as a case study of moral reasoning, but also as it connects to moral action.  

A further possibility is suggested by studies examining judgments about people who experience conflict (or not) before committing an immoral action (Starmans & Bloom, 2016). While adults, unsurprisingly, judge that when two people both act immorally, someone who is torn between doing the right thing and the wrong thing is morally superior to someone who wholeheartedly wants to commit the same immoral action. Three- to 5-year-olds, however, were much more likely than adults to judge that a person who had no internal struggle before acting immorally was morally superior. Coupled with children’s preference for the person who did not experience inner conflict before doing a moral action, these findings suggest that young children may have a preference for a “harmonious” self, regardless of whether this harmony is directed toward good or bad actions. Thus, intriguing questions remain about whether preschoolers truly prefer internal harmony, even if that harmony is directed toward bad actions.

Conclusions and Future Directions  

Taken together, the research reviewed in this article supports the idea that preschool-aged children exhibit robust intuitions about the nature of the self. They share adult intuitions that the self is...
fundamentally distinct from the body, though located in a particular location within the head, near the eyes. They also have an intuitive theory of personal identity that is based on spatiotemporal history and the continuation of certain psychological traits like memory, knowledge, and moral behaviors. Children are able to reason about conflict within the self, and undergo interesting developmental change in their judgments of the inner conflicts of others.

The research reviewed earlier also points to several important areas for future research. In particular, many of the studies reviewed here describe third-person attributions of a self, rather than first-person experiences of having (or being) a self. Do children think differently about the self when considering their own self versus the selves of others? Furthermore, how do children’s developing folk theories about the nature of the self connect to, or extend from, children’s phenomenological sense of their own self, which begins to emerge much earlier in infancy?

Another area ripe for future research is how children think about the selves of other nonhuman entities. The research reviewed focuses, reasonably enough, on how children think about humans. However, certain aspects of self may also be attributed to nonhuman entities such as animals, robots, or even animate shapes and puppets. How do children reason about these more minimal sorts of selves? Understanding which aspects of the self children attribute to these other entities not only is informative about children’s reasoning about nonhuman entities themselves, but is also essential for understanding the social and moral consequences associated with the attribution of particular aspects of the self.

Finally, the research described earlier suggests that children have stable, early-emerging intuitions about the self that are often strikingly similar to adults’ intuitions about the self. However, there is much more to be learned about the origins of these intuitions, and the extent to which they arise from core aspects of children’s early reasoning about agents, are a by-product of children’s normative cognitive maturation, and to what extent they are molded by experience.

While many interesting questions yet remain about the precise nature of children’s intuitions about the nature, persistence, and unity of the self, this discussion has outlined important areas for further investigation into our intuitive conception of the self and its properties, while providing evidence that empirical investigations of the origin of children’s theories of the self can offer valuable insight into deep philosophical questions, as well as providing a complementary perspective on the development of children’s developing self-views and the emerging distinction between self and other.

References


Descartes, R. (1667). Discourse on method, optics, geometry, and meteorology, trans (P. J. Olscamp, Trans.). Indianapolis, IN: Bobbs-Merrill. (Original work published 1637)


Flavell, J., Shipstead, S., & Croft, K. (1980). What young children think you see when their eyes are closed.


Supporting Information

Additional supporting information may be found in the online version of this article at the publisher’s website:

Data S1. References