Perspectives on Perspective Taking: How Children Think About the Minds of Others

S.A.J. Birch*,1, V. Li*, T. Haddock*, S.E. Ghrear*, P. Brosseau-Liard†, A. Baimel*, M. Whyte*

*University of British Columbia, Vancouver, BC, Canada
†École de Psychologie, Faculté des Sciences Sociales, Université d’Ottawa, Ottawa, ON, Canada
1Corresponding author: e-mail address: sbirch@psych.ubc.ca

Abstract

Perspective taking, or “theory of mind,” involves reasoning about the mental states of others (e.g., their intentions, desires, knowledge, beliefs) and is called upon in virtually every aspect of human interaction. Our goals in writing this chapter were to provide an overview of (a) the research questions developmental psychologists ask to shed light on how children think about the inner workings of the mind, and (b) why such research is invaluable in understanding human nature and our ability to interact with, and learn from, one another. We begin with a brief review of early research in this field that culminated in the so-called litmus test for a theory of mind (i.e., false-belief tasks). Next, we describe research with infants and young children that created a puzzle for many researchers, and briefly mention an intriguing approach researchers have used to attempt to “solve” this puzzle. We then turn to research examining children’s understanding of a much broader range of mental states (beyond false beliefs). We briefly discuss the value of studying individual differences by highlighting their important...
implications for social well-being and ways to improve perspective taking. Next, we review work illustrating the value of capitalizing on children’s proclivity for selective social learning to reveal their understanding of others’ mental states. We close by highlighting one line of research that we believe will be an especially fruitful avenue for future research and serves to emphasize the complex interplay between our perspective-taking abilities and other cognitive processes.

1. A MULTIPURPOSE TOOL: THE MANY FUNCTIONS OF MENTAL STATE REASONING

Humans’ capacity to reason about the unobservable inner workings of others’ minds has fascinated philosophers and scientists for thousands of years. The wealth of interest in this capacity is not at all surprising when you consider that virtually all human behavior is driven by underlying mental states. As such, to make sense of another person’s actions it is rarely sufficient to rely exclusively on observable aspects of the person’s external environment. Instead, one must also make inferences about one’s internal mental states, including, for example, his or her goals, intentions, desires, knowledge, beliefs, and emotional states. This concept is nicely illustrated in the kinds of stories developmental psychologists give to young children to see if they understand this important aspect of human behavior. Consider the following scenario: Henry has lost his pet rabbit. He hears a noise inside the shed in his backyard. He opens the shed door to find the neighbor’s dog and tears begin to stream down his face. To make sense of Henry’s actions of going to the shed, opening the door, and crying, it is necessary to make inferences about his mental states (italicized below for emphasis). For instance, one can infer that having lost his rabbit (a pet he presumably likes) he is motivated (desires) to find his rabbit. Having heard something in the shed he thought or believed his rabbit could be inside. Having found the neighbor’s dog instead he learned (now knows) his belief was false, his desire to find his rabbit thwarted, and the tears streaming down his face were because he was sad as a result of his unfulfilled desire. Impressively, even children as young as 2 years of age understand that a boy who is looking for his rabbit will be sad if he found a dog, whereas a boy who wanted to find a dog would be happy (Wellman, 1990). That is, even young children seem to understand

\footnote{Of course there are simpler explanations one could generate to account for Henry’s behavior, such as “he is afraid of dogs or hates dogs” and, even though they might not be as accurate, they still involve positing mental states such as fear or hatred.}
that it is not the outcome, or the physical reality of the situation, that matters most, but rather the person’s mental states, such as their goals and desires toward that reality, that ultimately explain human behavior. In fact, the field of Psychology, the study of the mind and human behavior, rests upon this fundamental principle.

Of primary interest in this chapter, however, is not the formal discipline of Psychology but rather what has been called “naïve psychology”: lay people’s (i.e., nonpsychologists, typical children, and adults) capacity to explain and predict human behavior by inferring and reasoning about their mental states. This capacity has received a variety of names in the literature such as folk psychology, common sense psychology, mind-reading, mentalizing, mental state attribution, perspective taking, role-taking, and perhaps most commonly “theory of mind.” For simplicity, we primarily use the terms theory of mind or perspective taking throughout, whereby we mean the processes involved in inferring, and reasoning about, the mental states of others. To be clear, this use of the term should be distinguished from its early use as an all or none concept that others possess minds (i.e., an understanding that people have mental representations that differ from reality), as in “Do chimpanzees (or young children) have ‘a theory of mind’?” (Premack & Woodruf, 1978; Wimmer & Perner, 1983).

In contrast, like many before us, we use the term more broadly to refer to the processes involved in reasoning about the specific contents of those minds, the particular thoughts, desires, or beliefs a person holds (not just the presence of a mind). Specifically, we contend that these processes (a) are partially innate, (b) develop over time and can be honed through experience, and (c) lie on a continuum with some individuals being better than others. Our goals in writing this chapter were to provide a skeletal review of the research questions developmental psychologists ask to shed light on children’s understanding of the inner workings of the mind, and most importantly, why such research is so incredibly valuable in understanding human nature and our ability to interact with, and learn from, one another.

As the “lost rabbit” scenario illustrates, our theory of mind or perspective-taking abilities are vital to understanding human behavior. They do not just help us make sense of behavior in retrospect (as in the “lost rabbit” scenario) but they also assist us in making inferences about how someone will behave in the future. For example, if you know that Adam is unaware that the baseball game was canceled at the last minute, you can anticipate that he will still show up for the game. But the many functions of theory of mind do not stop there. For instance, our theory of mind also allows
us to manipulate others, for better or for worse (e.g., we could tell Adam the
game was canceled, even if it was not, so that he does not show up; or we
could tell Vivian her Mom was not able to come to the game so she would be
pleasantly surprised when she did). We also regularly deploy our theory of
mind, particularly our ability to reason about what others will likely know
and not know, to communicate effectively with others. Whether we are
having a conversation with just one person, writing a manuscript, or giving
a lecture, we routinely make inferences about what information is likely
common knowledge to our audience (and does not need further explana-
tion) and what information needs to be elaborated upon. Finally, we also
use our theory of mind when learning from others. We do not passively
absorb any and all information others provide, but are selective in our learn-
ing. Both adults and children use a variety of cues to make inferences about
whether the information they encounter is from a knowledgeable source
(e.g., who is knowledgeable and under what conditions).

In this chapter, we begin with an overview of the earliest research in the
field of theory of mind that culminated in the birth of the classic false-belief
tasks (often described as the best “litmus” tests for a theory of mind). Next,
we describe research with infants and young children that created a puzzle
for many developmental psychologists, and briefly mention an intriguing
approach used by researchers to attempt to “solve” this puzzle. We then turn
to the wealth of research on children’s understanding of mental states that
goes beyond false-belief reasoning to highlight the rich and complex set
of processes involved in making inferences about the mental activities of
others. We give a brief nod to the value of studying individual differences
by highlighting their important implications for social functioning and
social well-being and indicate ways to foster better perspective-taking abil-
ities. Next, we review work illustrating the value of capitalizing on
children’s proclivity for selective social learning to shed light on their under-
standing of the mind. Finally, we close with a discussion of some open
questions in the literature that we believe are especially fruitful avenues
for future research and highlight the complex interplay between our
perspective-taking abilities and other cognitive processes.

2. A HISTORICAL OVERVIEW OF RESEARCH LEADING
TO THE BIRTH OF THE FALSE-BELIEF TASK

The question of how we develop the ability to reason about the
mental states of others dates back to the beginning of developmental psy-
chology. Often referred to as the founding father of cognitive development,
Jean Piaget proposed that young children are so fixated on their own point of view that they neglect other people’s perspectives, making it difficult for them to effectively communicate or cooperate with others. In his famous three-dimensional “three-mountains” test of children’s visual perspective-taking, it was not until 7 years of age that children could correctly identify another individual’s visual perspective of the mountains; prior to that children routinely selected their own view of the mountains (Piaget & Inhelder, 1956). Piaget argued that young children’s egocentrism prevents them from shifting from their point of view to reason about perspectives other than their own. For much of the 1960s and 1970s, the development of perspective taking was often viewed as the gradual decrease of egocentrism with age, despite challenges from Vygotsky and others who emphasized that children are very interested in the people around them and surprisingly attuned to them (Vygotsky, 1967; see also Flavell, Botkin, Fry, Wright, & Jarvis, 1968; Selman, 1971).

As research progressed, Piaget’s late-onset approach to children’s perspective taking, and ability to cooperate, began to lose support. An abundance of research showed that infants have a range of social abilities, including an inclination to cooperate with others (Trevathan & Hulbey, 1978), a capacity for joint attention (Scaife & Bruner, 1975), and a proclivity for imitation (Meltzoff, 1976; Meltzoff & Moore, 1977)—all abilities that require one to adopt perspectives different from their own. These findings suggested that infants are more in tune with other people than Piaget had surmised, and in some cases, were used to argue that infants are capable of making basic mental state inferences (e.g., infants realize that gaze indicates interest; Scaife & Bruner, 1975).

Around the same time, researchers took an interest in utilizing comparative psychology to better understand theory of mind. Vygotsky suggested that language and thought form a unique theory of mind understanding among humans that is not shared with other species that lack language. Yet, in what would become a landmark paper, Premack and Woodruf (1978) put forth a study suggesting that chimpanzees do understand that others have minds that govern behavior. Specifically, they examined a chimpanzee’s inferences about the goals of another individual by showing a chimpanzee (Sarah) a series of clips of a human agent struggling to complete a goal (e.g., reaching for a banana). Subsequently, Sarah was prompted to indicate what the agent would do next by examining photographs. Interestingly, Sarah was able to indicate what the agent would do to complete the goal (e.g., use a stick to reach for the bananas). From Sarah’s responses, the researchers suggested that chimpanzees can represent the goal.
of another agent, and predict what he or she will do next. However, in a subsequent set of commentaries, several researchers criticized their interpretation by suggesting that the chimpanzee’s performance could be explained by general problem solving abilities, and did not necessitate mental state reasoning (Dennett, 1978).

Critics argued that Sarah was simply indicating what she would do next in that context (e.g., grab a stick to reach the bananas). Philosopher Dan Dennett suggested that to truly examine whether chimpanzees can reason about others’ mental states, they must be able to predict an agent’s action even when it contradicts what the chimpanzee would do him or herself. That is, if a chimpanzee can predict an action that he or she would not do (e.g., because it is inconsistent with reality), then the chimpanzee cannot simply be projecting his or her own actions but truly considering another individual’s belief, or mental representation of reality. Dennett argued that to be certain that an individual can reason about mental states, one must show that they understand that mental states (or mental representations) can conflict with, or misrepresent, reality—as in the case of a false belief.

In response to this challenge, Wimmer and Perner (1983) gave birth to what is now often referred to as a classic false-belief task (also known as the “Maxi Task,” the “unexpected transfer task,” or the “Sally–Anne Task”). This task (see Fig. 1 for a depiction) was designed to examine whether

Fig. 1 A depiction of the Sally–Anne Task of false-belief reasoning (also referred to as the “unexpected transfer task” or “the Maxi Task”) first developed by Wimmer and Perner (1983). In a typical variant of this task, children hear a short story about two characters, e.g., Sally and Anne (A). The story describes a scenario where the protagonist (e.g., Sally) leaves an object in one location (location X) and subsequently leaves the scene (B and C). Then, in the protagonist’s absence, the second character moves the object from location X to location Y (D) and children are asked where the protagonist will look for the object upon her or his return (E). This figure consists of original artwork courtesy of Carrie Cheung.
human children understand that people can hold false beliefs. Here, children viewed a series of sketches outlining a scenario where a protagonist leaves an object in one location (location X) and subsequently leaves the scene. Then, in the protagonist’s absence, the object is transferred from location X to location Y and children were asked where the protagonist would look for the object upon her or his return.

Using Dennett’s logic, if children could predict that the protagonist would look for the object in location X, then they must understand that the mind can misrepresent reality. However, if children predicted that the protagonist would look for the object in location Y, so the argument went, then they did not understand that concept. Wimmer and Perner’s results revealed that 3-year-old children inaccurately indicated that the protagonist would look for the object in its current location (location Y). Between 4 and 5 years of age, a little more than half responded accurately and by 6–9 years of age, they all responded accurately (see also Baron-Cohen, Leslie, & Frith, 1985).

This general pattern of results, where young preschoolers often fail to reason about false beliefs whereas older preschoolers pass, was subsequently observed in other experiments testing children’s understanding of false beliefs, such as the Appearance–Reality Task (Gopnik & Astington, 1988) and the “Smarties” or “Unexpected Contents” Tasks (Perner, Leekam, & Wimmer, 1987). In an Appearance–Reality Task for instance, children are shown an object that looks like one thing (e.g., a rock) but is actually another thing (e.g., a sponge). After discovering the object’s true identity (e.g., by touching it), children are asked what another person will think the object is. That is, children are asked to reason about the false belief of another individual who should mistake the object for something else because of its appearance (e.g., falsely believe the sponge is a rock; Gopnik & Astington, 1988). Impressively, the aforementioned pattern of results holds across nearly 200 different experiments using these three types of false-belief tasks (see Wellman, Cross, & Watson, 2001 meta-analyses and ensuing commentaries).

3. WHAT CAN WE INFER FROM THE RESULTS OF THE CLASSIC FALSE-BELIEF TASKS?

Why 3-year-old children consistently fail the classic false-belief tasks has been, and still is, a topic of great debate (Rubio-Fernández & Geurts, 2013; Wellman et al., 2001 and subsequent commentaries). Although
several accounts of these findings have been put forth, two overarching views emerged. According to one view, sometimes referred to as the Conceptual Change view, 3-year-olds do not realize that the mind can misrepresent reality, but somewhere between ages 3 and 5 they experience a qualitative shift in their conceptual understanding of the mind (Perner, 1991; Wellman, 1990; Wellman et al., 2001). Another overarching view, sometimes referred to as a Processing Demands account, suggests that children cannot pass the task at 3 years of age because they have not fully developed other cognitive abilities required to pass the false-belief tasks (e.g., working memory, language, inhibitory control). According to this latter view, 5-year-olds’ ability to pass these tasks could reflect the maturation of their general cognitive capacities rather than a qualitative conceptual change in their understanding of the mind (Bloom & German, 2000; Fodor, 1992; Zaitchik, 1990). For instance, Birch and Bloom (2003) emphasized that classic false-belief tasks pose the unnecessary demand of requiring children to ignore their own specific knowledge (see also Bernstein, Atance, Loftus, & Meltzoff, 2004; Royzman, Cassidy, & Baron, 2003).

In the Sally–Anne Task, for instance, children are not only told that Ann moved the object from Location A (a requirement to test their understanding of false beliefs) but they are also told exactly where the object gets moved. The latter is not required to test one’s understanding of false beliefs, thereby making the task harder than it needs to be. This argument draws from an abundance of research with children and adults that shows that specific knowledge of an event, or fact, can bias one’s ability to reason about a more naïve perspective (for reviews, see Birch & Bernstein, 2007; Ghrear, Birch, & Bernstein, 2016; Hawkins & Hastie, 1990; Pohl, Bender, & Lachmann, 2002). This bias, the tendency to be swayed by one’s (current) knowledge when attempting to reason about a more naïve perspective, is referred to as the “curse of knowledge” (Birch & Bloom, 2003; Camerer, Loewenstein, & Weber, 1989). Importantly, this perspective-taking error occurs whether one is reasoning about someone else’s perspective or one’s own earlier perspective, although the latter manifestation is also referred to as “hindsight bias” or the “knew-it-all-along effect” (Bernstein et al., 2004; Fischhoff, 1977; Taylor, Esbensen, & Bennett, 1994; Sutherland & Cimpian, 2015; for a meta-analyses of 122 studies with adults see Christensen-Szalanski & Willham, 1991).
According to Birch and Bloom (2003), participants in the classic false-belief tasks not only have to make inferences about another individual’s false belief, but they also have to overcome the curse of knowledge. Importantly, research suggests that overcoming the curse of knowledge is particularly challenging for young children compared to older children and adults (Bernstein et al., 2004; Epley, Morewedge, & Keysar, 2004; Lagattuta, Sayfan, & Blattman, 2010; Lagattuta, Sayfan, & Harvey, 2014; Mitchell & Taylor, 1999). That is, younger children are more likely to overattribute their own knowledge than older children when considering the perspectives of others. For instance, Birch and Bloom (2003) found a significant decrease in the bias between ages 3 and 5.

Taylor et al. (1994) observed a similar decline in the “knew-it-all-along” manifestation of the curse of knowledge between ages 4 and 5. They found that when 4-year-olds learned new information (e.g., the color chartreuse) they claimed they knew it all along and were unable to differentiate between knowledge that they had known for some time (e.g., the color red), and knowledge they acquired that day (see also Sutherland, Cimpian, Leslie, & Gelman, 2015). Interestingly, young children also seem to think that their peers, and even babies, will know the information they just learned (Taylor, Cartwright, & Bowden, 1991).

Critically, the curse of knowledge bias has been shown to influence even adults’ false-belief performance—and few would question whether adults have a concept of false beliefs. Birch and Bloom (2007) presented adults with a variant of the classic false-belief task that used four containers instead of two. Participants read a story about Vicki who left her violin in the blue container. In Vicki’s absence, the violin was moved to a different container. In a Knowledgeable (cursed) condition participants were told exactly where the violin was moved (e.g., “it was moved to the red container”), akin to the classic false-belief tasks. In the Ignorant condition, participants were not told where it was moved (e.g., “it was moved to another container”). Then, participants were asked, “What are the chances Vicki will first look for her violin in each of the containers?” (in percentages). Importantly, participants in both conditions needed to realize that Vicki would hold a false belief (i.e., she would falsely believe that the violin was still in the blue container where she left it). However, when participants knew the exact location where the violin was moved they were more likely (compared to those in the Ignorant condition) to predict she would look where they knew it was, and less likely to predict she would look in
the false-belief location. That is, adults’ specific outcome knowledge can interfere with their ability to predict that another person will act according to his or her false belief.\(^b\)

Applying this same logic to understand the role the curse of knowledge plays in children’s false-belief reasoning, Ghrear, Haddock, Li, and Birch (2016) found that 3- and 4-year-old children performed significantly better on false-belief tasks when they were not required to overcome the curse of knowledge (with the effect driven predominantly by 3-year-olds). Specifically, 3- and 4-year-old children were presented with four stories where a protagonist (e.g., Sally) hides an object in one of four containers (e.g., blue container). Then, in the protagonist’s absence, another character placed the object in a different container (e.g., *When Sally was gone, Ryan hid Sally’s ball in a different spot! He may have hid it here, or here, or here*). Half of the time children were told exactly which container (e.g., “we know he hid it here”), the rest of the time they were not (e.g., “we do not know where he hid it”). After each story, children were asked to predict where the protagonist would look for the object. Children were more likely to accurately infer the protagonist’s false belief when they were not given specific information about where the object was moved, compared to when they were told exactly where the object was moved (the latter being most akin to the classic false-belief tasks\(^c\)). Although these findings should be replicated across a range of different false-belief tasks, they suggest that the classic false-belief tasks are unnecessarily difficult for younger children, raising the question of how false-belief reasoning develops (in-and-of itself) without the added burden of the curse of knowledge. Or put another way, it is unclear how much of the earlier scholarship showing age-related changes in these tasks stemmed from age-related changes in the curse of knowledge rather than false-belief reasoning, per se.

To be clear, the aforementioned results do not preclude the possibility that children also undergo some conceptual change in their understanding of the mind during the preschool period. Importantly, we wish to emphasize that the curse of knowledge is believed to be an inherent limitation on perspective taking (a by-product of an otherwise adaptive learning mechanism; see Hoffrage, Hertwig, & Gigerenzer, 2000) that not only contributes to

\(^b\) It is worth pointing out that in this study, among others, the curse of knowledge only occurs, or at least occurs to a greater degree, when the outcome makes sense or the participant can generate a plausible explanation for why it might be foreseeable to others (e.g., Pohl et al., 2002; Yopchick & Kim, 2012).

\(^c\) Many real-world situations involving false-belief reasoning may also require overcoming the curse of knowledge.
children’s difficulties with false-belief reasoning but also plays a significant role in a wide array of social perspective-taking “tasks” in children’s and adults’ everyday lives. As such, it warrants further investigation (see Section 8 for a discussion).

4. FALSE-BELIEF REASONING IN THE FIRST 2 YEARS OF LIFE?

In the past decade, a growing body of research emerged that further challenged the conceptual change view of the classic false-belief task. This research relying on an entirely different dependent variable (i.e., participants’ looking time) suggested that even infants can reason about false beliefs (Onishi & Baillargeon, 2005; Scott & Baillargeon, 2009; Surian, Caldi, & Sperber, 2007). Specifically, researchers argued that infants have a conceptual understanding of how a protagonist should behave when he or she has a false belief. This understanding was investigated using a violation-of-expectation paradigm, wherein infants were shown an event and two opposing “outcomes” for the event. One outcome was meant to fit the infants’ expectations and the other was meant to be incongruent with their expectations (if they understood false beliefs). Applying the logic that infants will look longer at an outcome that is inconsistent with their expectations (for a review, see Gweon & Saxe, 2013), the researchers measured participants’ looking times (i.e., how long they looked) at each outcome. For example, Onishi and Baillargeon (2005) examined 15- to 18-month-old infants’ looking time while observing a scenario where a protagonist leaves her toy in a hiding location and subsequently either leaves the scene or stays and observes. After a pause, the toy moves from one location to another. In one outcome the protagonist reaches for the toy at the original hiding location; in the other outcome she reaches to the new location. Critically, when infants viewed the scenario where the protagonist left before the toy moved, infants expected that the protagonist would behave according to a false belief. That is, infants looked longer when the protagonist reached for the new location rather than the original location where she had last seen it. The opposite pattern was observed when infants viewed the scenario where the protagonist stayed and observed the object move to the new location.

Again, how to interpret these data became a matter of great controversy. Some researchers suggested these data did not reflect false-belief reasoning but only that infants have certain expectations of behavioral rules
(De Bruin & Newen, 2012; Perner & Ruffman, 2005; Ruffman & Perner, 2005; Ruffman, Taumoepeau, & Perkins, 2012) or result from domain-general processes such as “low-level novelty” (Heyes, 2014). Moreover, violation-of-expectation tasks do not require participants to predict one’s actions based on a false belief, but instead only require that they make sense of those actions in retrospect (in contrast to classic false-belief tasks). These criticisms opened the door for research designs to examine early false-belief reasoning without relying strictly on looking time data or retrospective analyses. For example, Buttelmann, Carpenter, and Tomasello (2009) tested toddlers’ false-belief reasoning by capitalizing on their precocious helping tendencies (Warneken & Tomasello, 2006, 2007); here, one experimenter hid a toy inside one of two boxes and either stayed in the room or left. Then a second experimenter engaged the child and sneakily moved the toy from one box to the other and locked it. When the first experimenter returned, he tried (but failed) to open one of the boxes. The second experimenter encouraged the child to help. Wisely, instead of helping him open the box he was trying to open, children as young as 18 months appeared to recognize his goal (and his false belief about its location) and consistently helped the experimenter open the correct box.

These new types of designs that placed the participant in a much more active role harkened back to much earlier scholarship that similarly questioned the conceptual change interpretation of the false-belief tasks introduced in the 1980s (see Chandler & Birch, 2010 for review). For instance, Chandler, Fritz, and Hala (1989) demonstrated that 2.5-year-olds are capable of deceiving others, which appears to entail some appreciation that people can misrepresent reality (or hold false beliefs) as well as an understanding that the mental states of others can be manipulated. In their study, children were asked to help a puppet, named “Tony,” hide a treasure from an experimenter. The problem was that when Tony moved the treasure, it left behind a trail that would lead the treasure-hunting experimenter right to its actual location. The trails, however, could be wiped away, which is something the children learned and practiced earlier in the study. Thus, when children were asked to help Tony hide the treasure, they had the opportunity to alter the physical state of the world (e.g., by removing the trails) in order to manipulate the mental state of the treasure-hunting experimenter. Their results demonstrated that 50% of their youngest participants (2-year-olds) were already capable of deceiving the experimenter and went so far as to create new false trails to false treasure locations.
The research suggesting that infants and young children (younger than those who passed the classic false-belief tasks) could reason about false beliefs posed an interesting puzzle for developmental psychologists. If the interpretation of these findings were correct and young children could understand false beliefs under some circumstances (or using some measures), then why did older children fail the classic false-belief tasks? To reconcile this apparent disparity some researchers have proposed a dual processing, or two-systems, account (Apperly & Butterfill, 2009; Low, 2010; Low, Apperly, Butterfill, & Rakoczy, 2016; Low & Perner, 2012; Penn & Povinelli, 2007; Sabbagh, Benson, & Kuhlmeier, 2013) of theory of mind. According to these two-systems accounts, humans have one evolutionarily “old” system that processes social information (such as eye-gaze direction) implicitly. This system processes information relatively quickly and effortlessly but has clear limitations and is somewhat inflexible. In contrast, humans have a second, slower but more flexible, system that involves explicit reasoning or deliberation. Apperly and Butterfill (2009) propose that the explicit system comes with the development of language and higher-order executive functions (i.e., cognitive processes including attentional control, inhibitory control, working memory, and cognitive flexibility, as well as reasoning, problem solving, and planning) and argue that infants and nonhuman animals are not yet capable of System 2 perspective taking, but may succeed in false-belief paradigms that do not necessitate explicit reasoning, via System 1.

Consistent with this approach, Rhodes and Brandone (2014) tested 3-year-old children’s false-belief performance using both verbal and action-based responses. In their design, an experimenter (E1) left the room through one of two doors (curtains) and explained that they would ring the doorbell when they wanted to come back. On the way out, E1 placed a toy into a box in front of the door they were leaving through. Children completed two false-belief trials where another experimenter (E2) would engage the child and sneakily move the toy to the box in front of the other door. When the doorbell rang, the child was encouraged to open the door for E1; which door they opened served as a test of their action-based false-belief performance. To compare the child’s action-based responses against their verbal responses, on one of the two test trials, the child was asked an explicit verbal question “where does E1 think the toy is?” The results revealed that 3-year-olds’ actions suggested they understood another individual’s false belief (by opening the correct door); however, they were not able to correctly answer the explicit question by pointing or verbally stating where E1 would think the toy was. By completing action-based and explicit measures
in the same trial they were able to show the aforementioned discrepancy in children’s responses within the same task (but see He, Bolz, & Baillargeon, 2012 for data suggesting that it is not specifically verbal demands that account for the differences in performance, and an alternative explanation for differences that emerge between what they refer to as “spontaneous” vs “elicited” measures).

These new theoretical accounts offer appealing solutions to the puzzle regarding the discrepancy between early false-belief competence and the later emerging ability to pass the classic false-belief tasks. At minimum they offer a novel approach to the study of theory of mind that goes far beyond their applications to false-belief understanding. They raise a multitude of questions for future research, such as: When and why, in the course of evolution, did “System 2” emerge? When and why, in the course of development, does System 2 emerge? How do these systems develop? How do they interact (e.g., Does System 2 build upon System 1)? Do they have different neuroanatomical loci? Are both systems domain-specific? What is the relationship between System 2 and metacognitive processes outside the domain of theory of mind? If they are not the result of two different systems, exactly what makes performance on “elicited” tasks different from performance on more spontaneous measures (see He et al., 2012)? In short, we still have a lot to learn about the processes involved in reasoning about others’ mental states and how those processes develop. Consequently, there is much to be gained by comparing and contrasting the conditions under which children succeed and those under which they fail. However, to gain a comprehensive understanding of the processes involved in theory of mind and how they develop we must also “think outside the false-belief box.”

5. THINKING OUTSIDE THE FALSE-BELIEF BOX: THEORY OF MIND IS MUCH, MUCH MORE THAN REASONING ABOUT FALSE BELIEFS

One needs only to consider the complexities of their own mental worlds to realize that when it comes to reasoning about another’s mental activity there is much, much more than an understanding of false beliefs. Although the ontogeny of an explicit understanding of false beliefs (e.g., their ability to pass the classic false-belief tasks around age 4 or 5) has sometimes been considered the hallmark of children’s “theory of mind,” we suspect few would disagree that the parts of social cognition that children grasp earlier than 4, as well as those they grasp later, are just as important to our understanding of children’s conceptions of the mind.
Well before children pass the classic false-belief tasks, they routinely make inferences about the mental worlds of other people. In a simple but clever study, for instance, Repacholi and Gopnik (1997) provided evidence that by 18 months children can differentiate their own preferences and desires from those of another person. In their study, an experimenter indicated that she preferred broccoli over crackers (i.e., a preference that differed from the child’s preference for crackers). Afterward, the experimenter would hold his or her hand out as an indication of a desire for one of the snacks. The 18-month-olds tended to provide the experimenter with broccoli at rates above that expected by chance, but the 14-month-olds routinely offered the experimenter their preferred snack of crackers. Interestingly, Wellman and Liu (2004) demonstrated that children’s later understanding of false beliefs is predicated by this earlier understanding that others have desires that can be different from one’s own. Differentiating one’s own desires from someone else’s might seem like simply a necessary precondition; however, it raises the important question of how children come to carve up the world in terms of different mental states in the first place? Moreover, it highlights the importance of identifying exactly what processes are involved in making inferences about the various different kinds of mental states, such as goals, intentions, desires, emotions, knowledge, and false beliefs, not to mention other mental activity, such as forgetting, idea generation, speculating, and other mentalistic concepts, including bias, pretense, inference, opinion, and interpretation, to name just a few.

A full account of research that examines children’s understanding of the full array of mentalistic concepts is clearly beyond the scope of the current chapter (see Harris, 2006 for an excellent review of more of this work). Instead, we briefly review a few examples to highlight the scope and richness of the mental activities children must understand in order to function in the complex social world in which we live. The consequences of a failure to understand mental activity are perhaps best exemplified by research with children diagnosed with an Autism Spectrum Disorder (ASD). Indeed, the pervasive impairments that individuals along the spectrum display in mental state reasoning have lent support to the notion of an early developing and specialized system for mental state reasoning (Baron-Cohen, 1997). There is ever-mounting evidence to suggest that children with Autism display a relatively specific deficit in mental state reasoning and its precursors, such as gaze-following (Baron-Cohen et al., 1985; Baron-Cohen, Tager-Flusberg, & Cohen, 1993; Perner, Frith, Leslie, & Leekam, 1989). Children with ASDs, even in the absence of general cognitive deficits, fail to reason about false beliefs (Baron-Cohen et al., 1985),
and they fail to pay attention to subtle social cues, such as gaze direction that typically developing children use as indicators of others’ mental states from an early age (Baron-Cohen, Campbell, Karmiloff-Smith, Grant, & Walker, 1995). Moreover, children with Autism have difficulty appreciating whether an action is accidental or deliberate (D’Entremont & Yazbek, 2007) unlike typically developing children (Dunfield & Kuhlmeier, 2010). Children diagnosed with ASDs tend to engage in virtually no spontaneous pretend (or fantasy) play (Sigman & Ungerer, 1981), they struggle with sarcasm (Happé, 1993), and they have difficulties with other pragmatic aspects of language (for a review, see Baron-Cohen, 1997).

According to some accounts, this specialized system responds to specific inputs and can be “tricked” into activation with cues that normally cooccur with minds (e.g., eyes and/or socially contingent behaviors). Such cues can lead even 12-month-old infants to treat inanimate objects as if they had mental states (Johnson, 2003; Kuhlmeier, Wynn, & Bloom, 2003). Even 5-month-old infants appear to recognize (presumably at an implicit level) that there is a fundamental distinction between the principles that govern inanimate objects (i.e., naïve physics) and the principles that govern human behavior (i.e., naïve psychology) (Kuhlmeier, Bloom, & Wynn, 2004; Spelke, Phillips, & Woodward, 1995). This specialized system for reasoning about animate objects is perhaps best exemplified by research that shows young children can infer that humans, but not inanimate mechanical agents, are goal directed. For example, Meltzoff (1995) demonstrated that 18-month-old children will infer the intended actions of a human target, and subsequently complete the intended outcome, but do not do so when a mechanical arm performs the same action as the adult human. This suggests that from an early age children treat humans and their mental states as “special” sources of information (see also Woodward, 1998).

With age children become increasingly able to apply their developing mental state reasoning skills (presumably in conjunction with other cognitive skills) for ever more complex purposes. For instance, although young children are capable of deception under at least some circumstances (Chandler et al., 1989 reviewed earlier), this does not mean they are master manipulators. Peskin (1992), for instance, demonstrated that children have some serious problems deceiving others, even when the potential for rewards are high. In this study, 3- to 5-year-old children were presented with stickers of varying desirability and asked to choose which ones they liked most and least. Afterward, children were introduced to a puppet that they were told would always take the sticker that the child wanted most.
Thus, for this malevolent puppet to not end up with the sticker the child wanted, he or she would have to lie about their true preferences and indicate a false preference for an unwanted sticker. Three-year-olds in this study persistently revealed their true preferences across repeated trials. Four- and 5-year-olds, however, were increasingly able to conceal their preferences to obtain their most desired stickers. These results are consistent with more recent scholarship that suggests that although 3-year-olds are capable of detecting the foundations of malicious intent (i.e., an individual’s previous malevolence), it is not until ages 4 or 5 that children come to understand, and capitalize on, explicit references to deception (Mascaro & Sperber, 2009). Together, this work suggests that understanding that others can intend to deceive, figuring out how to avoid falling prey to those deceptions, and being able to manipulate others’ mental states in order to deceive them are all important pieces in the development of theory of mind.

With development children’s ability to reason about the mental states of others becomes increasingly complex and flexible (for reviews, see Astington & Baird, 2005; Doherty, 2009). For instance, Carpendale and Chandler (1996) argue that around age 7 or 8 children become aware that the mind interprets information and that two people can view the exact same thing and come to a different interpretation, sometimes referred to as an “interpretive theory of mind.” Later developments also include the ability to appreciate ironic statements (Filippova & Astington, 2008), the ability to recognize how different kinds of knowledge are acquired through different senses (e.g., touch vs taste; see Robinson, 2000 for a review), and the ability to recognize that people’s interpretations can result from various biases, including those stemming from self-interest or self-enhancement (Mills & Keil, 2005; see also Elashi & Mills, 2015), close personal relationships (Mills & Grant, 2009; Mills & Keil, 2008), or group membership (Bigler & Liben, 1993). As just one further example, by around 9–11 years of age, children demonstrate an understanding of faux pas (i.e., situations where one person unintentionally upsets another due to a lack of knowledge about some aspect of the situation; Baron-Cohen, O’Riordan, Stone, Jones, & Plaisted, 1999; Stone, Baron-Cohen, & Knight, 1998). For example, a friend bought new “retro” curtains to decorate her home, and upon visiting for the first time, you comment that the curtains look like they need to be updated. Detecting and understanding faux pas situations nicely illustrates the complexities of theory of mind reasoning because the child must not only detect that one person has upset another, but also that the violation occurred as a result of the offender’s uninformed knowledge state,
and that the resulting insult was unintended; it might also entail appreciating that the offender will likely be embarrassed by her mistake and the friend may be quite forgiving, and ultimately even amused, given the nature of the misunderstanding.

In sum, the ability to reason about the contents of other minds becomes increasingly sophisticated and nuanced across development. This development does not appear to have a clear end state, and even a so-called mature system is still remarkably error prone. Even adults’ mental state reasoning is often inaccurate (Apperly, Warren, Andrews, Grant, & Todd, 2011; Birch, 2005; Royzman et al., 2003). Moreover, vast individual differences exist, in adults as well as in children, a subject we turn to next.

6. INDIVIDUAL DIFFERENCES IN THEORY OF MIND DEVELOPMENT, THEIR POSSIBLE ORIGINS, AND THE IMPLICATIONS FOR FOSTERING THEORY OF MIND

As reviewed throughout, several “milestones” mark the ontogeny of theory of mind. In addition to individual differences in the onset of these milestones, children (and adults) differ in the degree to which they can, and do, use theory of mind in their everyday lives (Cutting & Dunn, 1999; Repacholi & Slaughter, 2003). In other words, despite having the ability to explicitly reason about other minds, individuals can vary in the frequency with which they do so (sometimes related to their motivation to do so) and the accuracy with which they do so. As an extreme example of how individuals can vary along these dimensions, paranoid schizophrenia is argued to involve an active (noticing minds everywhere) but inaccurate (delusional) mental state reasoning system (Crespi & Badcock, 2008). Typically developing children’s theory of mind abilities can also be scored along these two dimensions: propensities to attribute mental states (Severson & Lemm, 2016) vs accuracy in attributing mental states (Epley et al., 2004).

The individual differences approach to measuring theory of mind capacities has been especially fruitful at predicting children’s social competencies (Walker, 2005). For instance, children who score higher on theory of mind measures (e.g., tasks that measure emotional understanding, perspective taking, and false-belief reasoning) are more likely to engage in prosocial behavior (Deković & Gerris, 1994; Denham, 1986; Lalonde & Chandler, 1995; Nelson & Crick, 1999; Watson, Nixon, Wilson, & Capage, 1999). A meta-analysis of 76 studies of children between the ages of 2 and 12 years
suggests that individual differences in theory of mind are positively, albeit modestly, predictive of propensities for helping, sharing, comforting, and coordinating with others to achieve shared goals ($r=0.19$; Imuta, Henry, Slaughter, Selcuk, & Ruffman, 2016). Moreover, links between better theory of mind performance and more positive peer relationships appear to be at least partially mediated by the relationship between early prosociality and peer acceptance (Caputi, Lecce, Pagnin, & Banerjee, 2012). Conversely, children who display poorer theory of mind performance are not as well regarded by their peers (Dunn, 1996; Slaughter, Dennis, & Pritchard, 2002) and are more likely to be involved in bullying behavior, either as bullies, victims, or “bully victims” (those who serve as both aggressor and victim; Renouf et al., 2010; Shakoor et al., 2012; Sutton, Smith, & Swettenham, 1999).

Interestingly, Banerjee, Watling, and Caputi (2011) found a bidirectional relationship between mental state understanding in the context of faux pas and peer rejection. In a longitudinal study of 210 children aged 5–6 and 8–9 at Time 1, who were followed up 1 year later and again 2 years later, revealed that earlier peer rejection impaired the acquisition of faux pas understanding. Additionally, difficulties in faux pas understanding in the older age group predicted increased peer rejection, highlighting the complex relations between mental state understanding and peer relationships.

Researchers study individual differences in theory of mind to illustrate the important function theory of mind serves in navigating our social worlds. In addition, researchers also want to account for what these differences may tell us about the processes involved in the development of theory of mind. This is likely to shed light on potential interventions to foster better perspective-taking abilities. Broadly speaking, research exploring individual differences in theory of mind has emphasized the contributions of other skills (e.g., language, memory, inhibitory control) and the importance of environmental scaffolding (e.g., mental state discourse, the role of siblings).

As just one example of the complex interplay between theory of mind and other cognitive skills, Hughes (1998) demonstrated a relationship between executive function and theory of mind performance, even when controlling for age and verbal ability. More specifically, there was a significant correlation between working memory and false-belief reasoning, and between inhibitory control, deception, and false-belief explanation tasks. Interestingly, these relationships can change with age and may interact in complex ways. One possible reason for individual differences in theory of
mind is that some individuals have better language, memory, or inhibitory abilities that facilitate theory of mind, but that is unlikely the whole story. Precisely how these cognitive abilities interact with theory of mind is a matter of great interest and debate (Carlson & Moses, 2001; Carlson, Moses, & Claxton, 2004; Davis & Pratt, 1995; Flynn, 2007; Hala, Hug, & Henderson, 2003; Hughes, 1998; Landry, Miller-Loncar, Smith, & Swank, 2002; Moses & Carlson, 2004; Müller, Lieberman-Finestone, Carpendale, Hammond, & Bibok, 2012; Russell, 1996; Zelazo, Carter, Reznick, & Frye, 1997). For example, in considering the strong correlation between language and false-belief understanding, Jenkins and Astington (1996) suggest that theory of mind and language do not develop independently of one another, and show how early language abilities are extremely important to a child’s social cognitive development. The opposite is also true; children’s early theory of mind abilities (e.g., inferring intentions) are often critical for understanding the intended referents of word labels, as well as many pragmatic aspects of language, which may go a long way in explaining why children with ASD also have problems with language (see Bloom, 2000). Thus, individual variability in executive function and language abilities appear to account for some of the variability in theory of mind. However, precisely how these cognitive processes interact in producing measurable differences in mental state reasoning remains an important avenue for future research.

In twin studies, the similar performance on theory of mind measures between identical and fraternal twins suggests that environmental input is important for theory of mind development (Hughes et al., 2005). Interestingly, heritability estimates gradually decrease throughout the first decade of life (Hughes & Devine, 2015), emphasizing the increasing influence of environmental factors on theory of mind development. More specifically, the quantity and quality of language present in the child’s environment (and specifically mental state language, e.g., terms like “want,” “think,” and “know”) are important contributors to theory of mind development (Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991). This might explain why deaf children whose deafness went undetected for some time and were exposed to fewer conversations and less mental state discourse (whether in spoken or signed language) lag behind their typically developing counterparts in false-belief understanding (see Peterson & Siegal, 2000 for a review). In contrast, children with older siblings (and presumably greater exposure to intersibling and parent–child discourse) develop theory of mind somewhat earlier than those without older siblings (Lewis, Freeman,
Other kinds of experience, or practice, also appear helpful in improving one’s social perspective taking abilities. For example, both the presence and richness of pretend play (imaginary friends, impersonation of imagined characters) positively predict later false-belief understanding (Taylor & Carlson, 1997; Youngblade & Dunn, 1995). The benefits of pretend play are compounded by social or interactive pretend play (e.g., with siblings or peers), arguably because it requires increasingly complex metarepresentational abilities (e.g., reasoning about how individuals can hold different beliefs or representations of the world at any given time, and practice “taking on the role” of different people or characters; Schwebel, Rosen, & Singer, 1999).

Beyond the child’s immediate environment, one’s broader sociocultural context appears to influence the developmental trajectory of mental state reasoning. For instance, American children come to understand that others can have diverse desires earlier than children from China (Wellman, Fang, Liu, Zhu, & Liu, 2006). Similar differences have been found between Australian and Iranian children (Shahaeian, Peterson, Slaughter, & Wellman, 2011). Research accounting for the origins of cultural differences in theory of mind development is still in its infancy and will surely be an important avenue of future research. However, recent work suggests that cultural differences in parent–child discourse and specifically the use of ostensive pedagogical cues in child–parent–object triangulation (e.g., eye gaze, pointing, verbal referents) may play an important role in accounting for some cultural differences in mental state reasoning (Little, Carver, & Legare, 2016).

Given the research supporting the importance of experiential factors, such as the quality and quantity of mental state language input on theory of mind development, it stands to reason that increasing children’s exposure to mental state discourse can help foster theory of mind abilities. Training studies with younger children suggest that participating in mental state conversations causally advances theory of mind abilities. Ornaghi, Brockmeier, and Gavazzi (2011) randomly assigned 3- and 4-year-old children to a 2-month training program in which children either played language games and read stories containing mental state language (or to a control group that used the same time to engage in free play). They found that the former group showed significant improvement in emotion and false-belief understanding as well as metacognitive vocabulary compared to pretest performance (see also Dunn & Brophy, 2005;
Nelson, 2005). Lohmann, Tomasello, and Meyer (2005) conducted appearance–reality tasks with 3-year-old children and had them experience a deception (i.e., showing them a ball of soap that appears to be a golf ball). These researchers found that children who then engaged in a conversation about the deception using mental state terms performed better on subsequent theory of mind tasks compared with children who merely watched a discussion concerning the deceptive object.

Aside from training children directly, training studies with their families show that increasing mental state discourse in the child’s environment can foster theory of mind. For instance, observations of parents of 2- and 6-year-old children showed that parents’ use of mental state terms (e.g., think, know, believe) predicted individual differences in children’s false-belief understanding 4 years later (Ensor, Devine, Marks, & Hughes, 2014; see also Meins et al., 2002). One experimental study showed that preschoolers who overheard characters in a video discussing another person’s mental states improved their false-belief performance (Gola, 2012), impressively demonstrating that mental state discourse, even when it is not directed at the child, can help draw children’s attention to others’ perspectives and improve performance.

It is important to note that not all children may benefit similarly from the kind of training paradigms described in the aforementioned studies. In a training study of 3.5-year-old children who previously failed a false-belief task, researchers found that individual differences in executive function strongly predicted the degree to which children benefitted from theory of mind training (Benson, Sabbagh, Carlson, & Zelazo, 2013). Further research is needed to ascertain not only what other kinds of interventions can foster theory of mind (and which are most helpful), but also how individual differences in other skills can bolster or attenuate the benefits of different training strategies.

7. HOW SELECTIVE SOCIAL LEARNING CAN REVEAL CHILDREN’S UNDERSTANDING OF THE MIND

It should be no surprise that humans possess an array of cognitive abilities that seem tailored for reasoning about other minds as doing so has played an integral role in humans’ success as a cultural species throughout our evolutionary history (Henrich, 2015). In addition to its usefulness for the interpretation of others’ behavior, theory of mind comes in handy when using others’ actions and words as information sources. Social learning
(i.e., learning from others) is an efficient, relatively low-cost strategy to acquire information. However, people are sometimes poor sources of information, providing misleading information on purpose (lies, jokes), by accident (mistakes, ignorance, or misunderstandings), and because they frequently offer their opinions not just facts. Thus, theory of mind can render social information gathering more efficient by allowing one to consider others’ knowledge, beliefs, and intentions, to gauge the reliability of the information they convey. A large body of research has demonstrated that children are selective in their social learning (see Chudek, Brosseau-Liard, Birch, & Henrich, 2013; Mills, 2013; Stephens, Suarez, & Koenig, 2015) and frequently capitalize on cues to others’ mental states to guide their learning. For example, older preschoolers consider others’ intentions, preferring to accept claims from a “helper” rather than a “tricky” individual (Vanderbilt, Liu, & Heyman, 2011). As another example, preschoolers are more likely to accept an individual’s claim about a novel object if that individual claims familiarity with the object rather than ignorance, suggesting that they take a speaker’s knowledge into account when deciding whether or not to trust his or her statements (Sabbagh & Baldwin, 2001).

The remainder of this section focuses on children’s understanding of an individual’s knowledge (or its counterpart, ignorance) in social learning. Knowledge can be conceptualized as a state or a trait: One may wonder whether a person possesses a specific piece of knowledge in a given situation (their knowledge state) or whether an individual is generally knowledgeable across a wide variety of situations (their trait knowledge). There is evidence that children use cues to both of these aspects of knowledge to facilitate social learning.

A valuable cue to knowledge as a state is a person’s perceptual access to relevant information. Though preschoolers have an imperfect understanding of the relation between the different senses and different types of “knowing,” they understand simple relations such as looking can lead to knowing—at least some types of knowing, such as knowing an object’s visual identity (Pillow, 1989; Pratt & Bryant, 1990; Robinson, Champion, & Mitchell, 1999). In some situations they also distinguish informative from noninformative perception. For instance, 3- and 4-year-olds can infer whether they should trust an adult’s claims about an object’s identity depending on the modality of the adult’s perceptual access: they are more likely to trust claims about texture if the adult has touched (but not seen) the object, and about color if the adult has seen (but not touched) the object (Whitcombe & Robinson, 2000). In the absence of information about an
individual’s perceptual access, children can use indirect cues to someone’s knowledge state, such as markers of confidence. Preschoolers prefer to learn from people who currently appear confident rather than hesitant in their actions or statements, both when confidence (or lack thereof) is expressed verbally (Jaswal & Malone, 2007; Moore, Bryant, & Furrow, 1989) and nonverbally (Birch, Akmal, & Frampton, 2010).

There are also enduring individual differences in knowledge that can hold across many situations. Some people know a lot more than others about a broad variety of topics (e.g., they may be smarter or more educated); people also have specific niches of expertise that often generalize across a particular domain of knowledge (e.g., music, math, biology, pop culture). These individual differences depend on a combination of factors including the person’s prior experience, interests, aptitudes, and so on, which may not be readily observable in a given situation. In the absence of other cues, one cue that is often readily available and can signal who has useful information is who other people are observing; indeed, children at least as young as 3 years of age take advantage of this cue to guide their learning (Chudek, Heller, Birch, & Henrich, 2012).

Another way to assess someone’s credibility as a source of information is to track, over time, the accuracy of the information provided by an individual. All else being equal, someone who has repeatedly offered accurate information in the past is a safer bet for providing accurate information in the future over someone who has a history of being inaccurate. Impressively, children at least as young as 3 years of age appear to spontaneously track the accuracy of individuals (Birch, Vauthier, & Bloom, 2008) and prefer to learn new information from individuals who have repeatedly provided accurate information over those who have demonstrated inaccuracy or ignorance (Jaswal & Neely, 2006; Koenig, Clément, & Harris, 2004; Koenig & Harris, 2005). Older preschoolers demonstrate further sophistication by simultaneously tracking accuracy in different domains (Sobel & Corriveau, 2010). Preschoolers also take advantage of other cues such as age (Fitneva, 2010) or occupation (Lutz & Keil, 2002) to infer what someone is likely to know.

Even younger children appear sensitive to some cues to others’ mental states when seeking information (for review, see Poulin-Dubois & Brosseau-Liard, 2016). At 18 months, toddlers prefer to learn from a previously reliable labeler, e.g., someone who correctly labels a bird, over someone who incorrectly labels a bird an “apple” (Brooker & Poulin-Dubois, 2013). By 24 months, children moderate their imitation of an adult’s actions based on the nonverbal confidence cues demonstrated by that adult.
In terms of knowledge as a trait, 14-month-olds are more likely to imitate an individual’s actions if that individual has previously demonstrated correct conventional actions (e.g., putting a shoe on one’s foot) rather than incorrect ones (e.g., putting a shoe on a different body part; Zmyj, Buttelmann, Carpenter, & Daum, 2010). Precursors to these abilities may even be present in the first year of life (Tummeltshammer, Wu, Sobel, & Kirkham, 2014).

Although infants and children’s selective learning preferences are now being used as a window into their understanding of the mind, there is still some controversy about the relation between some of the observed social learning preferences and theory of mind. Are children really engaging in mental state reasoning when deciding from whom to learn? Or could they be using more simplistic strategies (e.g., “stop listening to anyone who says weird things”)? Research has tried to answer this question in a few different ways. One way is to examine correlations between children’s theory of mind and their performance on selective learning tasks that appear to reflect some understanding, tacit or otherwise, of knowledge. If reasoning about knowledge does underlie children’s learning strategies on these tasks, then we would expect children’s theory of mind to predict more savvy selective learning.

A few studies have obtained moderate positive correlations between their preference to learn from a more knowledgeable individual and their performance on one or more theory of mind tasks (DiYanni & Kelemen, 2008; DiYanni, Nini, Rheel, & Livelli, 2012). In one such study by Brosseau-Liard, Penney, and Poulin-Dubois (2015), 3- and 4-year-olds’ performance on Wellman and Liu’s (2004) theory of mind scale correlated with a preference to learn new words from a previously accurate (rather than inaccurate) puppet, and as predicted no correlation was revealed between theory of mind task performance and their preference to learn from an individual demonstrating physical strength. This finding suggests that theory of mind may specifically predict selective social learning strategies involving reasoning about knowledge (as opposed to, for instance, a general tendency to be choosy about information sources).

Similarly, Lucas, Lewis, Pala, Wong, and Berridge (2013) investigated the impact of exposure to Turkish, a language that marks the source of one’s knowledge with obligatory evidential markers. Turkish preschoolers were superior to children of the same age from England and Hong Kong on both theory of mind and selective word learning from an accurate individual. Moreover, across all three countries theory of mind performance predicted
stronger selective learning from a previously accurate source. These studies suggest a potential relation between greater theory of mind skill and savvier social learning, but not all studies that have looked for such a correlation have found one (Pasquini, Corriveau, Koenig, & Harris, 2007). So far, the search for relations between savvy social learning and performance on theory of mind tasks has been mainly focused on 3- and 4-year-old children using only a few theory of mind tasks. More research is needed to fully understand the relation between theory of mind and selective social learning from the most knowledgeable sources (including, but not limited to, the use of tasks tapping other facets of theory of mind and the employment of experimental designs).

There are additional indications that children’s selective learning preferences may constitute a demonstration of their knowledge understanding. For instance, by late in their preschool years (around age 5), children do not blindly apply simple rules (e.g., looking leads to knowing; ignore what previously inaccurate people say) when deciding from whom to learn. Instead, they appear to consider the situational relevance of the cues at hand. This is especially noticeable when children are put in a situation where multiple cues conflict. Older preschoolers seem to favor cues that are more objectively informative, such as one’s prior accuracy, over less informative cues such as one’s tendency to make confident claims (Brosseau-Liard, Cassels, & Birch, 2014). They also appear to correctly apply cues to state knowledge and trait knowledge in the situations where they are most relevant. As an example, imagine that one wants to find out what is inside an unmarked box. If one person makes a claim about the box’s contents, that person’s perceptual access to the contents (i.e., did they look inside?) is crucial to evaluating the accuracy of that person’s claim; it does not matter how accurate that person has been in the past, if they have not seen inside the box they almost certainly will not know what the box contains.\(^d\)

Furthermore, a person’s knowledge of episodic information (e.g., one’s knowledge of where your mother left her sunglasses) is not helpful for predicting that person’s knowledge outside that situation—it is situation specific. Conversely, a person’s accuracy about generic information, such as their knowledge of common facts or the labels for objects, is likely to be a better indicator of his or her knowledge in other situations. Although these may seem like complex notions, some research suggests that preschool-aged

\(^d\) Barring, of course, situations where they were informed in some other way (e.g., someone told them) and were unable to infer the contents from the outward appearance of the box.
children recognize them. For example, preschoolers grant greater weight to a history of accuracy in providing semantic rather than episodic information (Stephens & Koenig, 2015), rely more heavily on information about a person’s perceptual access over that person’s prior accuracy in an episodic learning situation (Brosseau-Liard & Birch, 2011), and appreciate that someone with a history of accuracy is no more likely to know about situation–specific information than someone with a history of inaccuracy (Brosseau-Liard & Birch, 2010).

Overall, this body of work demonstrates the potential for capitalizing on children’s selective social learning preferences to answer questions about their understanding of the mental states of others. Simultaneously, it serves to highlight the important role theory of mind plays in social learning—the type of learning that is critical for the cumulative transmission of information from one generation to the next, and arguably (see Henrich, 2015) what makes us unique from all other species.

8. FUTURE DIRECTIONS: THE VALUE IN UNDERSTANDING AN INHERENT LIMITATION ON PERSPECTIVE TAKING AND THE MECHANISMS INVOLVED

Throughout this chapter we have highlighted important areas for future research. Rather than reiterate them here, we chose to emphasize one particular line of research that we believe is key for a comprehensive understanding of the processes involved in perspective taking or theory of mind. In addition, by discussing other cognitive processes believed to influence perspective taking, we emphasize that to understand the development of theory of mind and its deployment in our everyday lives, we must consider its complex interplay with other cognitive processes.

To understand how humans reason about the minds of others, one needs to recognize an inherent cognitive limitation on perspective taking, namely the so-called curse of knowledge. As noted earlier, the curse of knowledge refers to the tendency to be biased by one’s current knowledge state when attempting to reason about (or from) a more naïve perspective. This intrinsic limitation on perspective taking appears to stem from an otherwise adaptive learning mechanism: Our brains are geared toward acquiring knowledge, not ignoring it. Although humans are prone to forgetting, the act of intentionally “unknowing” something is inherently difficult, perhaps impossible (see Golding, Long, & MacLeod, 1994). When humans acquire new information,
this information gets embedded into our knowledge structures, or mental representations, and these representations get updated rendering old information, or outdated perspectives, less accessible (Henriksen & Kaplan, 2003; Hoffrage et al., 2000). This knowledge updating is extremely adaptive, in fact critical, for learning and allows us to keep track of the most current state of affairs in an ever-changing environment (Hoffrage et al., 2000). The downside (the so-called curse): this learning mechanism, that is otherwise so adaptive, consequently interferes with our ability to reason as if we did not know that information, leading to biased perspective taking. Not surprisingly, this constraint on perspective taking is present cross-culturally (Heine & Lehman, 1996; Pohl et al., 2002), though the magnitude of the curse of knowledge bias may vary depending on cultural experience (McNamara, 2016).

Generally speaking, the curse of knowledge clouds our ability to predict what others know, and consequently, sometimes how they will feel or behave. Furthermore, given how much of communication rests upon making inferences about what one’s audience knows (e.g., what information is common knowledge and what information needs to be explained) it can also interfere with our ability to communicate effectively (Hinds, 1999; Pinker, 2014). In fact, the consequences of this perspective-taking limitation have been demonstrated across a wide-range of disciplines including politics, behavioral economics, and education as well as in several applied settings, such as business, law, and medicine (for reviews, see Ghrear, Birch, et al., 2016; Hawkins & Hastie, 1990).

It is important to note, however, that this by-product of human learning cannot explain, in and of itself, why young children and aging adults tend to show the curse of knowledge bias to a greater degree than older children and young adults. To account for the U-shaped developmental trajectory in the magnitude of the bias (Bernstein, Erdfelder, Meltzoff, Peria, & Loftus, 2011) one needs to consider the developmental changes that occur in other cognitive processes that can reduce or accentuate its effects. To date, the specific cognitive processes that influence this bias are still a matter of great debate in the adult literature. According to an inhibitory control account, the bias can be partially overcome by inhibiting the contents of one’s knowledge (Bayen, Pohl, Erdfelder, & Auer, 2007; Groß & Bayen, 2015). According to this account, if you know that Trump won the election and you want to recall what you thought before the election (e.g., how probable you thought it was that Trump would win) you must attempt to inhibit your newfound knowledge of the outcome. The same logic applies if you are reasoning about another person’s perspective, rather than your own earlier perspective. In other
words, one’s ability to suppress the contents of one’s current knowledge state influences the magnitude of the curse of knowledge. Consistent with the age-related changes in the magnitude of the bias, inhibitory control improves over development and declines in old age (Groß & Bayen, 2015). However in two studies by Bernstein, Atance, Meltzoff, and Loftus (2007), children’s hindsight bias (or curse of knowledge) and theory of mind performance were correlated (i.e., the greater the bias, the worse one’s theory of mind performance), but inhibitory control did not mediate that relationship. Of course, future work along these lines may shed additional light on the role of inhibitory control in perspective taking, and the curse of knowledge bias in particular, as inhibition appears to play some role in the successful deployment of one’s theory of mind (as noted in Section 6).

A second mechanism that some have argued plays a role in the curse of knowledge is fluency misattribution (Harley, Carlsen, & Loftus, 2004). This account draws from a large body of research with adults that shows that people use (and sometimes misuse) the fluency with which information is processed (or comes to mind) to make a variety of perceptual, cognitive, and affective judgments (Bernstein, Whittlesea, & Loftus, 2002; Jacoby, Woloshyn, & Kelley, 1989; Sanna, Schwarz, & Small, 2002; Winkielman, Schwarz, Fazendeiro, & Reber, 2003; Zajonc, 1968). According to this account, the curse of knowledge is exaggerated when one misattributes the source of the “fluency” or ease with which the information comes to mind (or with which the information is processed). The logic goes like this: instead of recognizing that the ease with which the information came to mind is due to our prior exposure to the information, we misinterpret that ease as resulting from the objective ease, foreseeability, or prevalence of the information. For example, if you know who won a recent election the information tends to come to mind quickly and easily and you can mistake that ease for it being easier, or more widely known, than it really is. In contrast to an inhibitory control account, it is not the ability to suppress the content of one’s knowledge (who won the election) that influences the magnitude of the bias but rather one’s ability to recognize, and correctly attribute, the source of the fluency with which it came to mind. In fact, fluency misattribution may be one type of source monitoring error.

Source monitoring includes tracking, encoding, and recalling the source of one’s knowledge, such as how or when that knowledge was acquired (Johnson, Hashtroudi, & Lindsay, 1993). For example, source monitoring is involved in recalling how and when you learned that the
capital of Syria is Damascus (i.e., your memory of the source of that knowledge). The potential role of source monitoring in perspective taking is somewhat intuitive. If you can remember, for instance, where and when you learned the location of the “Space Needle” (e.g., when you visited Seattle in high school) you are probably better positioned to gauge the likelihood that others shared a similar experience to acquire that knowledge. On the other hand, if you cannot recall how you learned the meaning of the word “confounded” you may have more difficulty gauging how well known the meaning of that word is, compared to words whose source you do recall (e.g., loquacious, studying for Scholastic Aptitude Tests, or SATs). As mentioned previously, fluency misattribution appears to be a failure to recognize the source of one’s fluency (e.g., it is fluent because of your prior exposure rather than its objective frequency; Jacoby, Kelley, Brown, & Jasechko, 1989).

Young children are notoriously bad at source monitoring (Gopnik & Graf, 1988; Robinson, 2000). Substantial improvements in source monitoring occur across the preschool period in line with the age-related decline in the magnitude of the curse of knowledge (Birch & Bloom, 2003). Of course, to determine whether (or how) source monitoring, inhibition, and fluency misattribution contribute to the curse of knowledge one needs to go beyond correlations with age. Importantly, these processes may not be mutually exclusive but may work in tandem to ameliorate or compound the effects of the bias. They also may not represent the full set of processes that influence the curse of knowledge. We believe experimentally testing the roles of each of these mechanisms is a fruitful avenue for further research, especially given how pervasive the curse of knowledge bias is, and how profound its implications are for many real-world situations.

Ultimately, a better understanding of the development of theory of mind, and the processes that support it, will shed light on both the inner workings of the human mind and the ability to reason about those inner workings. This line of work is also key to identifying ways to improve perspective taking. This is an especially fruitful avenue of research considering the myriad of benefits associated with accurate perspective-taking abilities, including better social emotional health, greater academic success, and more satisfying relationships, to name just a few (see Repacholi & Slaughter, 2003 for review). To emphasize this point we quote Ickes (1997) referring to people with what he called “empathic accuracy” (i.e., accuracy at inferring the thoughts and feelings of others): “All else being equal, they are likely to be the most tactful advisors, the most diplomatic
officials, the most effective negotiators, the most electable politicians, the most productive salespersons, the most successful teachers, and the most insightful therapists.” (p. 2). His quote should also remind us that there is a myriad of work on perspective-taking abilities outside the field of developmental psychology (e.g., in social psychology, cognitive science, behavioral economics, philosophy, political science, education). Insights from those disciplines will prove valuable when researching the development of perspective-taking abilities (and vice versa). Developing techniques to improve perspective taking in early childhood may be particularly beneficial because early childhood is a time when cognitive malleability is high. Fortunately (or unfortunately as it always depends on your perspective), the mental lives of people are so incredibly rich and complex that there is no end in sight to our ability to hone our perspective-taking skills (or in the number of research questions we can ask).

ACKNOWLEDGMENTS

We would like to especially thank Janette B. Benson for her very helpful feedback on an earlier draft of this manuscript as well as the exceptionally helpful Research Assistants in the K.I.D. Studies Centre at the University of British Columbia who assisted with the literature reviews and editing the manuscript (in alphabetical order): Carrie Cheung, Gini Choi, Marisa Gagne, Nataly Kaufman, Parky Lau, Caitlin Mooney, and Janane Sri. A special thanks to Carrie Cheung for the artwork for Fig. 1.

REFERENCES


