Child Development, July/August 2007, Volume 78, Number 4, Pages 1374–1394

Hindsight Bias and Developing Theories of Mind

Daniel M. Bernstein
Kwantlen University College and University of Washington

Cristina Atance
University of Ottawa

Andrew N. Meltzoff and Geoffrey R. Loftus
University of Washington

Although *hindsight bias* (the ''I knew it all along'' phenomenon) has been documented in adults, its development has not been investigated. This is despite the fact that hindsight bias errors closely resemble the errors children make on theory of mind (ToM) tasks. Two main goals of the present work were to (a) create a battery of hindsight tasks for preschoolers, and (b) assess the relation between children's performance on these and ToM tasks. In two experiments involving 144 preschoolers, 3-, 4-, and 5-year olds exhibited strong hindsight bias. Performance on hindsight and ToM tasks was significantly correlated independent of age, language ability, and inhibitory control. These findings contribute to a more comprehensive account of perspective taking across the lifespan.

Hindsight bias is typically studied in cognitive and social psychology. It occurs when outcome knowledge influences the judgments we make for a naïve other or a naïve ''prior'' self. For instance, armed with the knowledge that New Orleans suffered a devastating flood, we are more apt to think that ''we knew it would happen all along.'' In hindsight bias one's present knowledge influences one's recollection of previous beliefs. Interestingly, having to ignore or override one's own current knowledge is a component of many tasks used to assess theory of mind (ToM) in young children. For example, in the classic false belief task involving a change in location, children must reason about a character who does not share their own knowledge. To pass this task, children must not be swayed unduly by their own knowledge about the actual state of the world. Typically, success on this task occurs between 4 and 5 years of age (e.g., Wellman, Cross, & Watson, 2001).

The concepts of hindsight bias and ToM are related: both involve perspective taking and the misattribution of knowledge to the past self or a current other. There has been some empirical and theoretical work suggesting that aspects of ToM development may be related to a general, lifelong cognitive bias wherein one's own current knowledge warps judgments about the beliefs of a naïve other or the prior self (Atance & Meltzoff, 2006; Bernstein, Atance, Loftus, & Meltzoff, 2004; Birch & Bloom, 2003, 2004; Epley, Morewedge, & Keysar, 2004; Jacobs & Klaczynski, 2002; Keysar, Lin, & Barr, 2003; Royzman, Cassidy & Baron, 2003; Taylor, 1988; Taylor, Esbensen, & Bennett, 1994). However, to date, there has not been a systematic study of whether and how hindsight bias and ToM relate. We will begin by discussing hindsight bias and ToM and the ways these constructs relate.

Hindsight Bias

Hindsight bias in adults has been documented in many domains, including legal decisions (Harley, 2007; Kamin & Rachlinski, 1995), medical diagnoses (Arkes, Wortman, Saville, & Harkness, 1981), consumer satisfaction (Zwick, Pieters, & Baumgartner, 1995), sporting events, and election outcomes (Leary, 1981, 1982). In each case, people armed with advance knowledge of an outcome overestimate the likelihood of that particular outcome, in essence claiming that they ''knew it all along'' (Wood, 1978).

Hindsight bias has been measured using two experimental designs: the hypothetical design and the memory design (Pohl, 2004). Both yield comparable levels of hindsight bias. In the hypothetical design,

This research was supported by grants from NSF (SBE-0354453) and NIH (MH41637). We thank Jennifer Amsterlaw, Stephanie Carlson, David Liu, Betty Repacholi, Jessica Sommerville, and three anonymous reviewers for helpful comments and suggestions. We thank Joy Durham, Pallavi Shukla and Deniz Tahirouglu for assistance with data collection and Ryan Godfrey and Craig Harris for technical assistance.

Correspondence concerning this article should be addressed to Daniel M. Bernstein, Institute for Learning and Brain Sciences, Box 357988, University of Washington, Seattle, WA 98195. Electronic mail may be sent to Daniel.Bernstein@kwantlen.ca.

which we used in Experiment 1, participants who know the outcome to a question must estimate what they themselves or a naïve peer knows or would have thought without the benefit of outcome knowledge. For instance, a participant receives the answer to a challenging question: the Statue of Liberty is 151 feet from base to torch. Next she is asked to estimate what a naïve other would say if asked this same question. The participant says, ''175 feet.'' In actuality, a naïve person would rarely be this accurate, and the research shows that if the participant is not given prior knowledge about the height of the Statue of Liberty, she does not attribute such an accurate answer to the naïve person.

In the memory design, which we used in Experiment 2, participants must recall their original answer to a problem after receiving feedback about the correct answer. Thus, suppose, that Participant 1 says that the Statue of Liberty is ''200 feet'' and Participant 2 says, ''100 feet.'' They then receive the correct answer (''151 feet'') and must recall their original answer. The research shows that both participants will distort their memory of their own past judgment and gravitate toward their current knowledge. Participant 1 will deflate her past judgment and might now say, ''175 feet,'' while Participant 2 will inflate his past judgment and might say, ''125 feet.'' Their memory of their previous answer seems to be warped (but not completely replaced) by the newly learned fact. Hindsight bias is robust, occurring across a wide range of time intervals (from minutes to years) between the initial judgment, the outcome information, and the second judgment. The bias even occurs after participants are explicitly warned to avoid it (Pohl & Hell, 1996).

What causes hindsight bias? One account posits that individuals automatically update their beliefs with new information, rendering the original information inaccessible (Fischhoff, 1975). More recent theories maintain that hindsight bias results from a biased reconstruction of the original memory trace, using the outcome as a cue. On this view, the outcome information coexists with the original memory trace rather than altering or overwriting it (Pohl, Eisenhauer, & Hardt, 2003; Hoffrage, Hertwig, & Gigerenzer, 2000; see Hawkins & Hastie, 1990, for review).

Despite the substantial literature on hindsight bias that exists for adults, there has been little work on the developmental origins and trajectory of hindsight bias (but see Bernstein et al., 2004; Birch & Bernstein, 2007; Pohl & Haracic, 2005). This stands in stark contrast to the domain of theory of mind in which research has focused on developmental issues.

Theory of Mind

The ToM literature chiefly concerns the child's understanding that others have mental states, such as beliefs, desires, and intentions. An important and extensively researched aspect of ToM is children's understanding of false belief (Wimmer & Perner, 1983). Children's difficulty appreciating that others can hold beliefs that differ from reality (i.e., a false belief) is manifested in the errors that young children make. Several tasks have been created to trace the development of false belief understanding over the preschool years. In the change in location task, children witness a character placing an object in one location and then leaving the room. In the character's absence, children observe the object being moved to a new location. Children are then asked where the character will look for his object upon his return. Of interest is whether children state that he will look for it where he presumably believes it to be (where he saw it placed last) or where it really is. Children younger than about 4 years of age typically state the latter option. In another task, the unexpected contents task, children are presented with a crayon box and asked what they think is inside (Gopnik & Astington, 1988; Perner, Leekam, & Wimmer, 1987). After they answer, ''crayons,'' the box is opened to reveal candles. The box is then closed and children are asked what they first thought was inside the box before it was opened. Three-year-olds typically state, ''candles,'' while older children correctly state, ''crayons.'' A recent meta-analysis demonstrated that this pattern of performance on false belief tasks is robust across different procedures and cultures (Wellman et al., 2001; Wellman & Liu, 2004). In addition to difficulties with false belief, children younger than 4 years of age also tend to fail tasks that require them to acknowledge that the way something appears to oneself or others does not always match its true identity (appearance-reality tasks) (Flavell, Green, & Flavell, 1986; Gopnik & Astington, 1988; Taylor, 1988; Taylor & Flavell, 1984).

Relation Between Hindsight Bias and Theory of Mind

Hindsight and ToM tasks share at least one crucial feature: In both tasks, a participant learns that her initial belief was incorrect and then must reason about the belief of her ''naïve'' prior self or a naïve other. There are, however, important differences between these tasks. First, in many ToM tasks, participants must realize that either they previously held, or that another person holds, a *false* belief (or in the case of the appearance-reality task, that appearances can be misleading)–that is, that belief and reality can differ.

Thus, children must understand the concept that the mind is capable of misrepresentation or multiple representations about the same reality. Whereas 3-year-olds seem to have difficulty with this, older children typically do not (e.g., Wellman et al., 2001; Wellman & Liu, 2004). In contrast, in the hindsight task, the requirement is not that participants recognize that the mind can misrepresent reality, but rather, they must strive to recapture their previous judgment in the face of new knowledge. Moreover, in the classic hindsight task, an individual rarely holds his initial belief with great certainty—there is usually only a very limited commitment to it. Thus, the term ''belief'' may not describe the participant's initial judgment in a hindsight bias test–the term ''hunch'' may in fact be more accurate. For example, one only has a hunch about how tall the Statue of Liberty is, not a firm belief about it. This difference in one's initial certainty about the state of the world may distinguish the two tasks in ways that are important to consider. We explore this possibility in more detail in the General Discussion.

These differences between hindsight and ToM tasks are not trivial and may underscore the observed differences in performance. That adults easily pass classic false belief tasks but show robust hindsight bias suggests that ToM and hindsight bias are differentiable. Despite this, we also suggest that the two are related. This is because an important requirement of these tasks is that one must not be swayed unduly by newfound knowledge in reasoning about one's own prior judgment or a judgment that will be made by a naïve other. Indeed, some have argued that ToM errors may be exacerbated by the fact that children and adults have trouble ignoring knowledge that they possess when trying to estimate what another, naïve person knows–a phenomenon previously labeled as ''realist bias,'' ''epistemic egocentrism,'' or ''the curse of knowledge,'' (Birch, 2005; Birch & Bloom, 2004; Camerer, Loewenstein, & Weber, 1989; Jacobs & Klaczynski, 2002; Mitchell, Robinson, Isaacs, & Nye, 1996; Nickerson, 1999; Royzman et al., 2003; Taylor, 1988). In this vein, one might predict that children who show less hindsight bias should be better able to suppress or set aside their own knowledge in the context of a false belief task, thus allowing them to consider that someone else (or their prior self) may hold a false belief about the situation at hand.

Related to this is the suggestion that difficulties with the executive function skill of inhibitory control–the ability to ''suppress potentially interfering thought processes or actions'' (Carlson, Moses, & Hix, 1998, p. 672)–lie at the heart of false belief tasks specifically and, more generally, of people's difficulty ignoring outcome information (Birch & Bloom, 2003; Birch, 2005; Carlson et al., 1998; Carlson, Moses, & Breton, 2002; Moses, 2001; Royzman et al., 2003). In a false belief task, the actual state of affairs may be quite salient to the young child. Carlson et al., (1998) argue that the child has trouble inhibiting reference to reality–perhaps even in the presence of some understanding of false belief. They argue that one reason that it is hard to inhibit referring to reality is that people strive to be accurate in their descriptions of the world. Such a tendency could affect both false belief and hindsight tasks, because, in both, one must resist describing the world as it actually is (i.e., accurately) in favor of what one believed it to be. Inhibitory control skills show substantial development during the preschool years (e.g., Carlson, 2005). Moreover, researchers have found a significant positive correlation between ToM and inhibitory control (Carlson & Moses, 2001; Carlson et al., 2002; Frye, Zelazo, & Palfai, 1995; Hughes, 1998; Scullin & Bonner, 2006), although the underlying causal explanation for this correlation is a matter of intense debate (e.g., Perner & Lang, 2000; Russell, 1996).

In summary, two abilities need to be considered when contemplating the relation between theory of mind and hindsight tasks: (a) understanding that the mind is capable of misrepresentation (e.g., the concept of false belief), and (b) recapturing one's previous judgment in the face of new knowledge about the current reality by, possibly, inhibiting/overriding this knowledge. Whereas both of these abilities pertain to performance on ToM tasks, only the second pertains to performance on standard hindsight tasks (adults understand that the mind can misrepresent reality). We know that ToM improves during the preschool years, and hindsight bias persists throughout life. So, how do these constructs empirically relate to one another? We explore this question in the present study.

Experiment 1

There has been no empirical work directly examining the relation between hindsight bias and ToM errors in young children. The main challenge in studying this relation is developing hindsight tasks that are simple enough for children to understand, yet sensitive enough to detect variation in their performance. Most hindsight tasks used in adult cognitive science contain complex counterfactual language that is difficult for young children to understand (e.g., what would you have said was the outcome to the 19th century war between the British and the Gurkas of Nepal if you did not know that the British won? (Fischhoff,

1975). One of our main goals in this paper was to develop a comprehensive battery of hindsight tasks suitable for preschool-aged children but that would also be appropriate for adults (thus ensuring that we are indeed measuring "hindsight bias"). We have previously developed one such task (Bernstein et al., 2004; Harley, Carlsen, & Loftus, 2004). In a baseline condition, child and adult participants identify degraded images of common objects as the images gradually clarify on a computer screen. Soon after, in a hindsight condition, participants learn the identity of each object *before* estimating when a naïve peer would identify it as it clarified. As expected, preschool children and adults showed robust hindsight bias by overestimating their naïve peer's knowledge (Bernstein et al., 2004).

In the present work, we (a) design additional hindsight tasks and use these to assess the development of hindsight bias over the preschool years, (b) test the relation between children's performance on these newly designed tasks and on classic ToM tasks, and (c) administer inhibitory control tasks to assess whether the relation between hindsight bias and ToM is mediated by inhibitory control skills. To achieve these goals, we administered two different hindsight tasks (Computer and Real Object), a battery of ToM tasks, two inhibitory control tasks, and a measure of general language ability. This last measure was included because significant relations have already been reported between language and ToM (e.g., Astington & Jenkins, 1999).

Research Questions

1. (a) Does the hindsight bias that children demonstrated on a previous computer hindsight task (Bernstein et al., 2004) generalize to a real object version? (b) Will the magnitude of hindsight bias change with age?
2. Does hindsight bias relate to ToM? If hindsight bias is connected to children's difficulty with false belief (Birch, 2005; Birch & Bloom, 2003; Royzman et al., 2003), then these measures should be inversely related. Thus, the more hindsight bias one displays, the worse one should perform on false belief measures.
3. Does hindsight bias relate to inhibitory control? If inhibitory control is implicated in people's difficulty ignoring privileged information, as has been suggested (Birch, 2005; Birch & Bloom, 2003; Royzman et al., 2003; see also Friedman & Leslie, 2004; Leslie, German, & Polizzi, 2005; Leslie, Friedman, & German, 2004; Moses, 2001), then inhibitory control should be negatively related to hindsight bias. Thus, the worse one performs on inhibitory control measures, the more hindsight bias one will demonstrate.

Method

Participants

Three groups of children participated: 3.5-year olds (M = 42.0 months, SD = 0; 12 female); 4.5-year olds (M = 54.0 months; SD = 0.46; 12 female); and 5.5-year olds (M = 66.3 months, SD = 1.23; 12 female). There were N = 24 children per group. Children came from the Seattle metropolitan area. Two 3-year-old participants failed to complete the study and were replaced. Sixty-four parents reported the race of their child: There were 58 White children, four Asian children, one American Indian/Alaska Native child, and one Black/African American child.

Materials

The tasks for Experiment 1 are listed in the top half of Table 1.

Computer hindsight. Materials, counterbalancing, and procedures resembled those used previously (Bernstein et al., 2004). Specifically, stimuli consisted of eight line drawings of common objects: airplane, bicycle, chair, clock, glasses, keys, scissors, and telephone. We scaled pictures of each object to fit within a 245 × 245 pixel square on a Macintosh G4 Power Book. We then degraded each object in two ways: (a) by adding pixel noise, and (b) by cropping. We chose these degradations for their ecological validity; for example, objects can be obscured by random noise when viewed through a dirty window, or cropped by an obstruction, such as the framing around a window. We also chose these degradations because they could be reproduced easily in the Real Object Hindsight task (described later). For the Pixel procedure (*Pixel Computer*), we changed a proportion of image pixels to random grayscale values. We characterized Degree of Pixel degradation as the proportion pixels changed. For the Crop procedure (*Crop Computer*), we cropped each object and presented it in an expanding fashion, starting in the middle and expanding to the borders (see Figure 1). We measured Crop in terms of distance (see Bernstein, Loftus, & Meltzoff, 2005, for full description of degradation procedures; but roughly, imagine looking at a scene through a window: the scene becomes progressively more cropped as one moves back further from the window).

For each object and for each degradation type, we created 30 increasingly degraded images so that the

Table 1
List of Tasks Used in Experiments

Hindsight bias	Theory of Mind	Inhibitory Control	Language	
		EXPERIMENT 1		
Pixel Computer	Unexpected Contents	Bear/Dragon	PPVT	
Crop Computer	Change in Location	Day/night		
Pixel Real Object	Appearance Reality			
Crop Real Object	Occluded Pictures			
Hindsight bias	Theory of Mind	Inhibitory Control	Language	Working Memory
		EXPERIMENT 2		
Real Object Pixel	Unexpected Contents	Card Sort	PPVT	Count and Label
Discrete	Change in Location	Day/Night		Backward Dig. Span
Continuous	Appearance Reality			
	Occluded Pictures			

Note. PPVT = Peabody Picture Vocabulary Test-III; Backward Dig. Span = Backward Digit Span.

differences between successive degraded images were roughly equal perceptually. Each object clarified from fully degraded to fully clear. However, we only displayed every third image (of the 30 increasingly degraded images) to provide distinct stopping points that would permit us to prompt children for a response (see Bernstein et al., 2004, Experiment 2).

A *trial* consisted of a single clarifying object along with the participant's associated response. The participant's task on half the trials (Baseline) was to identify the object as soon as possible. At the outset of each Baseline trial, participants were unaware of the object's identity. Participants' task on the remaining trials (Hindsight) was to view the same objects that they had identified in the Baseline condition and estimate when a same-age peer (the Sesame Street puppet, Ernie) could identify the objects. Ernie occupied a nearby plastic black box. Children met Ernie and were told that he was the same age as them (e.g., ''Ernie is four years old just like you''). Children also learned that when Ernie was inside his home, he could not hear or see anything outside his home. At the outset of each hindsight trial with Ernie, the object appeared in full clarity while Ernie remained in his home. Children named the object and were reminded that Ernie could not see it. The fully clarified object then disappeared from the screen, and the experimenter said, ''Tell me when Ernie can see that the picture is a [chair].'' The experimenter retrieved Ernie from his home and placed him directly in front of participants. The object then clarified as in the Baseline condition. Participants completed the two Baseline conditions before completing the two Hindsight conditions.

Participants sat facing a laptop computer, with an experimenter seated directly beside them. After each stopping point in the Baseline (and Hindsight) conditions, the experimenter asked participants: ''What does it look like to you (Ernie) now?'' The experimenter told children that if they did not know or that it did not look like anything, they could say, ''don't know'' or ''nothing.'' The experimenter typed children's responses. The entire procedure consisted of four blocks of four trials per block. For each block of trials, degradation type (Pixel, Crop) and outcome knowledge (Baseline, Hindsight) remained constant (i.e., no intermixing of conditions within a block). Object order remained constant across all participants. Regardless of the point at which participants correctly identified the object, the experimenter prompted participants to indicate what the object looked like to them (Baseline) or Ernie (Hindsight) at *all* 10 stopping points.

Real object hindsight. We used eight real objects, measuring up to 5 inches long and five inches high: blue rubber shoe, red car, baseball, brass horse, yellow sailboat, green dinosaur, black and white hairbrush, and green coffee cup. A yellow rubber duck served as a practice object. The experimenter placed each object on a platform that was at eye level to the participant. This platform stood inside a rectangular plastic box that sat on one of its long sides. A black piece of paper covered the back of the box. Ten separate laminated transparency sheets (for Pixel Real Object) or 10 black, heavy-bond paper sheets (for Crop Real Object) were placed in a three-ring binder, which sat atop the box. The sheets hung in front of the toy, thus obscuring its appearance. In other words, there was a real toy inside

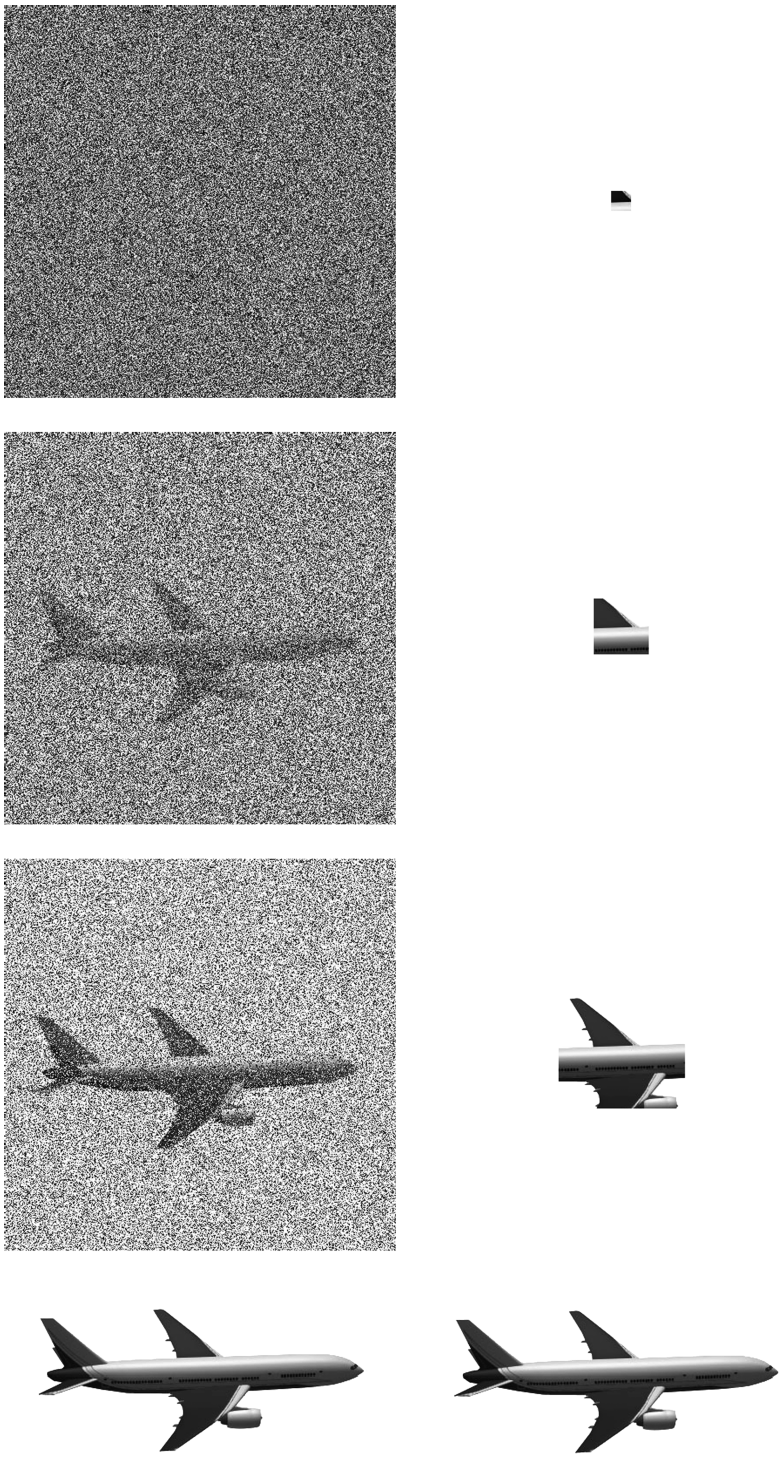

Figure 1. Examples of the Pixel (left-hand column) and crop (right-hand column) degradation procedures used in the Computer Hindsight task in Experiment 1.

a box, and we duplicated the visual pixilation and crop that were achieved on the computer screen. In the *Pixel Real Object*, each sheet contained a different laminated transparency. Each transparency contained a unique set of black dots that covered 10% of the surface area of the sheet. Note that objects are easily identified behind 10 blank transparencies. In the *Crop Real Object*, each sheet contained a different black

sheet of heavy-bond paper. Each sheet contained successively larger squares cut from the center, starting with a 3 × 3 mm square and increasing by 3 × 3 mm per square until the ninth sheet (27 × 27 mm square cut-out). The final black sheet contained a 126 × 200 mm rectangle cut from its center, permitting easy identification of the object. In both the Pixel Real Object and Crop Real Object tasks, the 10 sheets together occluded the object so that it was impossible to identify.

The experimenter told participants that they would now play the ''hide it game.'' As in the Computer Hindsight task, participants first completed the Baseline condition, where they attempted to identify each of the eight objects (i.e., four in pixel and four in crop) as it gradually improved in clarity. To accomplish this, the experimenter removed one sheet at a time and then asked participants: ''What does it look like to you now?'' When participants correctly identified the object, the experimenter recorded how many sheets had been turned. In the Hindsight condition, participants saw the object at the start of each trial. The experimenter then covered the object with all 10 filter sheets and reminded children that Ernie had not seen the object being placed in the hide-it box. The experimenter retrieved Ernie from his home and placed him directly in front of participants. Participants then saw each object clarify as it had in Baseline. After the experimenter removed each sheet, he asked participants, ''What does it look like to Ernie now?'' The experimenter recorded the point at which participants reported that Ernie could see the object, by recording the number of sheets that had been turned. As in the Computer Hindsight task, regardless of the point at which participants correctly identified the object, the experimenter prompted participants to indicate what the object looked like to them (Baseline) or Ernie (Hindsight) for all 10 sheets. Participants completed the two Baseline conditions before completing the two Hindsight conditions.

We characterized Degree of Pixel Real Object degradation for the Baseline and Hindsight conditions separately as the number of transparency sheets that had been turned. Note that the 1-to-1 correspondence between the percentage of unique black dots appearing on each sheet (10%) makes this characterization of degradation possible. We also measured Degree of Crop Real Object degradation for the Baseline and Hindsight conditions separately as the area of the square cut out corresponding to the number of sheets that had been turned.

We calculated hindsight bias for each hindsight task type (computer and real) and degradation type (pixel and crop) as the ratio of the mean identification point in the baseline condition divided by the mean identification point in the hindsight condition (Bernstein et al., 2004, 2005). For example, a subject who identified objects on average at the eighth filter in baseline and the sixth filter in hindsight would obtain a hindsight ratio of $8/6 = 1.33$. We then log transformed each hindsight ratio due to some extreme scores (± 2 SD from the mean).

Theory-of-mind tasks. We administered four tasks to index ToM: Unexpected Contents (e.g., Gopnik & Astington, 1988), Change in Location (e.g., Wimmer & Perner, 1983), Occluded Pictures (Gopnik & Astington, 1988; Taylor, 1988), and Appearance-Reality (Flavell, Flavell, & Green, 1983). For all but one of these tasks (Change in Location) participants answered two experimental questions, one about their own prior belief and the other about a naïve, same-age peer named Ellie. Ellie was a doll, located in a black box (in fact, Ellie's and Ernie's homes were identical and stacked atop one another). Participants met Ellie and learned that when she was inside her home, she could not hear or see anything outside her home. At the outset of each trial with Ellie, the experimenter reminded participants that Ellie was naïve (e.g., ''Ellie has never looked inside *this* box before'' for the Unexpected Contents task).

In the *Unexpected Contents* task, children were shown a crayon box, asked what they thought was inside and then shown that it contained candles. With the box closed, children were asked two test questions (Self: ''When you first saw this box, before we opened it, what did you think was inside?'' Other: ''What does Ellie think is inside?'' and a control question (''What is really inside the box?''). In the *Change in Location* task, children were introduced to two characters. One character places a ball in one location (a box) and then leaves the room. The other character then moves the ball to another location (a cupboard). The original character then returns to the room. Children were then asked one experimental question (''When Billy comes back inside to play, where will he look for the ball?'') and two control questions (''Where did Billy put the ball?'' and ''Where is the ball really?''). In the *Occluded Pictures* task, children were shown a book with three pictures. Only part of each picture was visible. Each visible part looks like animal ears–a fact that is confirmed when the full picture is revealed: bear ears in one case and cat ears in the other. For the third picture, what looks like bunny ears turns out to be a sunflower. The final picture is then covered up, and children are asked two test questions (Self: ''When I first showed you this window, all closed up like this, what did you think was underneath?''

Other: ''What does Ellie think is under the window?''), and a control question (''What is really under the window?''). In the *Appearance-Reality* task, children were shown what looks like a rock. They then discover that it is really a sponge. Children are then asked two experimental questions (Self: ''When you look at this with your eyes right now, what does it look like?'' Other: ''What does it look like to Ellie?'') and a control question (''What is it really and truly?''). Children only received credit for passing these questions if they answered the accompanying control question(s) correctly.

Inhibitory control tasks. We used two different Inhibitory Control (IC) tasks–Day/Night and Bear/Dragon–chosen from the battery used by Carlson et al. (2001) for their high reliability and correlations with false belief measures. Both of these tasks are categorized as ''conflict'' (as opposed to ''delay'') inhibition tasks, because the child is required to suppress his knowledge of reality while activating an alternate representation of it–this requirement being very similar to that of a false belief task (Carlson et al., 2002). As such, we reasoned that these tasks would be good candidates to assess a possible relation between IC, ToM, and hindsight bias. In the *Day/Night* task, children are instructed to say ''Day'' when they see a card with a picture of the moon and stars on it and ''Night'' whenever they see a card with a picture of a sun on it. This task consisted of two practice trials with corrective feedback, followed by 16 test trials. Children received 1 point for each correct response and 0 for each incorrect response. Children's first response was always scored, even in those instances in which there was a self-correction. We then calculated a proportion correct. In the *Bear/Dragon* task, a variant of ''Simon Says,'' children are told to follow the commands of a ''nice Bear'' puppet, while ignoring the commands of a ''naughty Dragon'' puppet (e.g., ''stick out your tongue;'' ''touch your feet''). The task consisted of five practice trials with corrective feedback, followed by 10 test trials (five Bear, five Dragon, in alternating order). Responses were scored from 0-3 for each trial. A full commanded movement = 3 when given by Bear, 0 when given by Dragon; a partial commanded movement = 2 for Bear, 1 for Dragon; a wrong/different movement = 1 for Bear, 2 for Dragon; and no movement = 0 for Bear, 3 for Dragon. Bear and Dragon scores were summed. We then calculated a proportion score out of 30.

Language ability. Children completed the Peabody Picture Vocabulary Test-III (PPVT-III; Dunn & Dunn, 1997), a measure of general language ability. Raw scores were computed by subtracting the number of errors from the ceiling item, as described by Dunn and Dunn. Age-standardized scores were used in all analyses.

Procedure

We tested children individually in one videotaped session (lasting approximately one hour). Parents observed the session on a TV monitor from an adjacent room. Each testing session began with three blocks of 12 PPVT-III trials. Testing ceased as soon as a child failed eight or more trials in a block. If the child passed the first three blocks, s/he completed three more blocks of PPVT-III trials in the middle of the session, and the remaining PPVT-III trials at the end of the session–until failing a block. The same male experimenter tested all children.

Counterbalancing

Upon completion of the three blocks of the PPVT-III, children either completed the Computer, or Real Object, Hindsight task (i.e., both baseline and hindsight trials) followed by the other hindsight task. We also counterbalanced degradation type in the hindsight tasks such that half the participants received Pixel before Crop on both the Computer and Real Object Hindsight and vice versa for the remaining participants. The ToM and Inhibitory Control tasks were interleaved in a counterbalanced order within the two hindsight tasks (e.g., children were given a false belief or inhibitory control task after completing two baseline or hindsight trials). Finally, we counterbalanced Self and Other questions in the ToM tasks.

Results and Discussion

We address each of our research questions in turn. For all tasks except the PPVT-III and Day/Night ($N = 71$ for each) we had complete data for all 72 participants. In cases of missing data, we used list-wise deletion.

Preliminary Analyses

Proportion correct on each of the ToM and inhibitory control tasks by age group appear in the top half of Table 2. Means and standard errors of each of the hindsight tasks by age group appear in Figure 2. As can be seen in Table 3, the ToM measures were interrelated (Cronbach's $\alpha = .85$) and so an overall measure of ToM performance was calculated by summing children's performance across each of the seven test questions (range = 0 to 7). The inhibitory control tasks were unrelated ($r = .07$), and therefore kept separate. The Hindsight tasks were interrelated, and therefore combined to form a Hindsight Bias scale (Cronbach's

Table 2
Proportion Correct (and standard error) on ToM, Inhibitory Control, and Working Memory Tasks by Age Group

Age	UC	CL	OP	AR	D/N	B/D
			EXPERIMENT 1			
3	.19 (.07)	.08 (.06)	.21 (.07)	.21 (.08)	.59 (.06)	.83 (.05)
4	.63 (.08)	.33 (.10)	.69 (.07)	.65 (.09)	.73 (.05)	.97 (.01)
5	.79 (.07)	.35 (.10)	.74 (.07)	.67 (.08)	.77 (.05)	.99 (.01)

Age	UC	CL	OP	AR	D/N	CS	CLA	BDS
				EXPERIMENT 2				
3	.25 (.07)	.25 (.09)	.33 (.08)	.59 (.09)	.61 (.07)	.56 (.08)	.39 (.10)	1.39 (.17)
4	.52 (.08)	.46 (.10)	.50 (.08)	.69 (.09)	.68 (.05)	.90 (.06)	.90 (.05)	2.37 (.21)
5	.96 (.03)	.71 (.09)	.85 (.06)	.98 (.02)	.81 (.04)	.99 (.01)	.98 (.02)	3.08 (.19)

Note. UC = Unexpected Contents; CL = Change in Location; OP = Occluded Pictures; AR = Appearance Reality; D/N = Day/Night; B/D = Bear/Dragon; CS = Card Sort; CLA = Count and Label; BDS = Backward Digit Span level (out of 5). Critical questions for self and other were averaged for Unexpected Contents, Occluded Pictures and Appearance Reality. Standard errors appear in parentheses.

$\alpha = .69$). As can be seen in the top half of Table 4, the significant correlations between age and ToM and between age and the Inhibitory Control measures reflect age-related changes similar to those reported in previous research (e.g., Carlson et al., 2001; Wellman et al., 2001).

Computer Versus Real Object Hindsight Bias

Computer Hindsight Bias and Real Object Hindsight Bias were significantly correlated ($r = .43$, $p < .001$); all three age groups showed hindsight bias on all four hindsight tasks (see Figure 2). In what follows, we list the mean hindsight bias for each of the four hindsight tasks and for each of the three age groups (Mean bias ± 95% Confidence Interval for 3-, 4-, and 5-year olds, respectively. Note that mean bias values ± the 95% Confidence Interval that fall above 0 are significant): Pixel Real Object (.34 ± .14; .21 ± .13; .25 ± .13); Pixel Computer (.61 ± .44; .65 ± .40; .70 ± .42); Crop Real Object (.79 ± .36; .54 ± .32; .58 ± .33); Crop Computer (.31 ± .20; .46 ± .23; .49 ± .22). Collapsing across degradation type, the magnitude of hindsight bias was similar for Real Object Hindsight ($M = .45$, $SEM = .06$) and Computer Hindsight tasks ($M = .42$, $SEM = .05$), $t < 1.0$ and remained stable with age ($F < 1.00$ for both comparisons). Thus, both hindsight measures yielded significant and similar amounts of hindsight bias in all three age groups.

Relation Between ToM and Hindsight Bias

ToM correlated significantly with Hindsight Bias ($r = -.31$, $p = .008$, see top half of Table 4). The negative correlation indicates that the worse one performed on ToM tasks, the more hindsight bias one showed. To test more conservatively the relation between hindsight bias and ToM, we conducted a hierarchical multiple regression by entering age, then language ability, then inhibitory control (Day/Night, Bear/Dragon), and finally hindsight bias to predict ToM performance. Age (R^2 Change = .29, $p < .001$) and language (R^2 Change = .08, $p = .005$) accounted for significant variance in ToM performance, whereas inhibitory control (R^2 Change = .05, $p = .069$) accounted for marginally significant variance in ToM performance after controlling for age and language ability. Hindsight bias accounted for an additional 9% of the variance in ToM when entered last into the regression equation (R^2 Change = .09, $p = .001$). All four predictor variables together accounted for 51% of the variance in ToM performance. We reran this regression on each age group separately, and the pattern remained. That is, hindsight bias accounted for 5% to 27% unique variance in ToM after controlling for language ability and inhibitory control. Finally, we ran a regression equation in which we entered all predictor variables in one step. The resulting model had three predictor variables accounting for significant variance in ToM: Age ($\beta = .33$, $p = .002$), Hindsight bias ($\beta = -.31$, $p = .001$), and Bear/Dragon ($\beta = .27$, $p = .01$). In sum, hindsight bias accounted for significant variance in ToM performance, even after controlling for age, language ability, and inhibitory control.

Relation Between Inhibitory Control and Hindsight Bias

As can be seen in the top half of Table 4, both Day/Night and Bear/Dragon correlated with ToM ($r = .36$ and .39, respectively), replicating previous work (Carlson et al., 2001). Neither of the inhibitory control

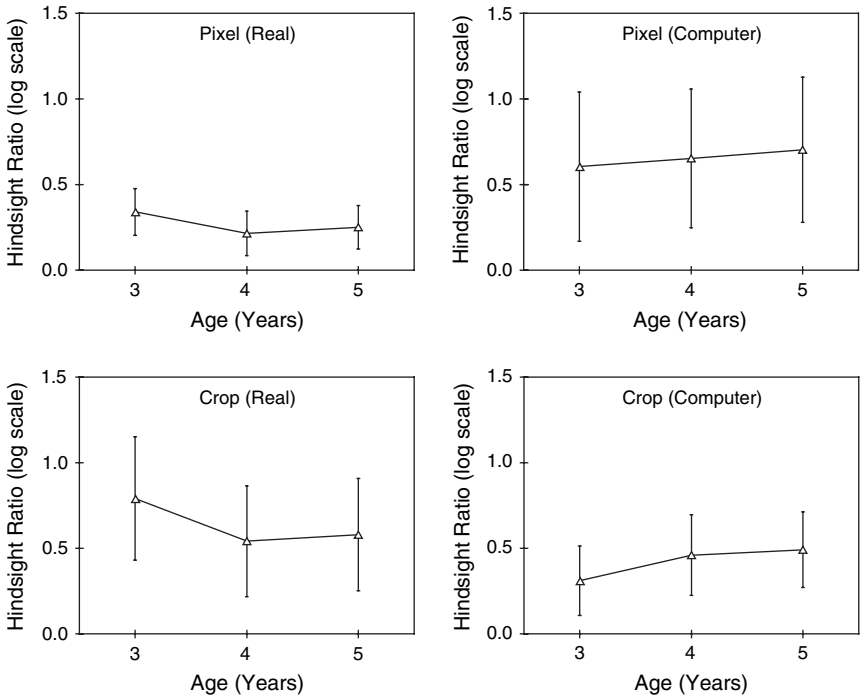

Figure 2. Magnitude of hindsight bias as a function of age in each of the four hindsight tasks in Experiment 1 (Pixel Real Object, Pixel Computer, Crop Real Object, Crop Computer). Values greater than 0 indicate hindsight bias. Error bars indicate 95% confidence intervals.

tasks correlated with hindsight bias: Day/Night and hindsight bias (r = -.06); Bear/Dragon and hindsight bias (r = .06). Thus, although both inhibitory control and hindsight bias related to ToM, inhibitory control and hindsight bias were unrelated.

These data demonstrate that (a) preschoolers exhibit robust hindsight bias by claiming that a naïve peer can identify objects on a computer or behind a series of filter screens at a more degraded level than they themselves could; (b) hindsight bias correlates significantly with performance on classic ToM tasks, even after controlling for age, language ability, and inhibitory control; and (c) hindsight bias does not correlate with inhibitory control.

The findings from Experiment 1 show that preschoolers have difficulty ignoring what they now

Table 3
Zero-Order Correlations for ToM, Inhibitory Control and Hindsight Tasks in Experiment 1

	CL	OP	AR	D/N	B/D	PRO HB	CRO HB	PC HB	CC HB
UC	.11	.67*	.71*	.28*	.25*	−.54*	−.23ˆ	−.25*	−.15
CL		.25ˆ	.48*	.05	.30*	−.23ˆ	−.12	−.09	−.07
OP			.64*	.26*	.46*	−.48*	−.32*	−.28*	−.24ˆ
AR				.36*	.50*	−.57*	−.28ˆ	−.35*	−.25ˆ
D/N					.07	−.19	−.08	.00	.00
B/D						−.19	−.11	.20ˆ	.14
PRO HB							.43*	.52*	.44*
CRO HB								.30*	.29*
PC HB									.73*

Note. UC = Unexpected Contents; CL = Change in Location; OP = Occluded Pictures; AR = Appearance Reality; D/N = Day/Night; B/D = Bear/Dragon; PRO HB = Pixel Real Object Hindsight Bias; CRO HB = Crop Real Object Hindsight Bias; PC HB = Pixel Computer Hindsight Bias. CC HB: Crop Computer Hindsight Bias. Critical questions for self and other were averaged for Unexpected Contents, Occluded Pictures and Appearance Reality.
*$p < .05$.
ˆ$p < .10$.

Table 4
Zero-Order (and Partial) Correlations

Task	Age	Language	Day/Night	B/D	HB
			EXPERIMENT 1		
ToM	.51*	.40*	.36* (.14)	.39* (.23ˆ)	−.31* (−.35*)
Age		.39*	.27*	.42*	−.005
Language			.27*	.39*	−.14
Day/Night				.07 (−.12)	−.06 (−.02)
B/Drag					.06 (.12)

Task	Age	Language	Day/Night	Card Sort	WMemory	PRO HB	DIS and CONT HB
				EXPERIMENT 2			
ToM	.64*	.45*	.17 (−.05)	.41* (.01)	.67* (.32*)	−.35* (−.29*)	−.39* (−.33*)
Age		.29*	.32*	.51*	.66*	−.21ˆ	−.19
Language			.12	.41*	.54*	−.10	−.20ˆ
Day/Night				.12 (−.06)	.19 (−.05)	−.31* (−.26*)	−.16 (−.10)
Card Sort					.54* (.19)	−.01 (.14)	−.12 (.03)
WMemory						−.13 (.03)	−.10 (.12)
PRO HB							.24* (.20)

Note. ToM = Theory of Mind; Language = Peabody Picture Vocabulary Test (PPVT-III), standardized score; B/D = Bear/Dragon; HB (Experiment 1) = average hindsight bias of the Pixel Real Object, Crop Real Object, Pixel Computer, and Crop Computer Hindsight tasks; WMemory = Working Memory: average standardized scores of Count and Label and Backwards Digit Span; PRO HB = Pixel Real Object Hindsight Bias; DIS and CONT HB = Discrete and Continuous Hindsight Bias, average standardized scores of Discrete and Continuous Hindsight Bias. Partial correlations controlled for age and language ability.
*$p < .05$.
ˆ$p < .10$.

know when trying to estimate the beliefs of themselves or others in a naïve state and, thus, similar to adults, show robust hindsight bias. We did not detect any age-related changes in hindsight bias. Inhibitory control and hindsight bias were unrelated. This result is inconsistent with the claim that inhibitory control mediates the relation between Hindsight Bias and ToM (Birch, 2005; Birch & Bloom, 2003; Royzman et al., 2003; see also Leslie, German, & Polizzi, 2005). Hindsight bias correlated modestly with ToM, a finding that is consistent with the idea that both types of tasks share common features. We return to this point in the General Discussion.

Experiment 2

In Experiment 2, we sought to replicate and extend the results of Experiment 1. We replaced one of our inhibitory control measures (Bear/Dragon) with the Dimensional Change Card Sort (Frye et al., 1995; Zelazo, Mueller, Frye, & Marcovitch, 2003), permitting us to retest the mediating influence of inhibitory control on the link between hindsight bias and ToM. We retained the Real Object Hindsight task from Experiment 1 and also developed several new hindsight tasks using a different hindsight bias experimental design: the memory design. This permitted us to further explore the developmental trajectory of hindsight bias and to again test the relation between hindsight bias and ToM. In addition to these new tasks, we added two measures of working memory to provide a more complete battery of executive function. Others have shown that working memory relates to ToM (e.g., Gordon & Olson, 1998; Keenan, 1998). Adding working memory tasks also permitted a more stringent test of the correlation between hindsight bias and ToM. If the latter two constructs are directly related, then their relation should persist even after controlling for working memory ability.

Research Questions

1. Does the hindsight bias that children demonstrated on our Real Object hindsight task in Experiment 1 replicate and generalize to yet other hindsight tasks?
2. Does hindsight bias correlate with ToM, even after controlling for age, language ability, inhibitory control, and working memory?
3. Do hindsight bias, working memory, and inhibitory control tasks relate? Based on the results of

previous research and theories in the literature, inhibitory control and working memory should correlate. However, based on the results of Experiment 1, inhibitory control and hindsight bias should not correlate. We made no firm prediction about the relation between working memory and hindsight bias.

Method

Participants

Three groups of 24 children participated: 3.5-year olds ($M = 42.2$ months, $SD = 0.51$; 12 female); 4.5-year olds ($M = 54.6$ months; $SD = 0.58$; 12 female); 5.5-year olds ($M = 66.3$ months, $SD = 0.55$; 12 female). Children came from the Seattle metropolitan area as in Experiment 1 and represented families of varying socioeconomic status and ethnic backgrounds. One 3-year-old participant failed to complete the study and was replaced. Fifty-four parents reported the race of their child: There were 44 White children, seven Asian children, one American Indian/Alaska Native child, and two children with multiple races.

Materials

The tasks for Experiment 2 are listed in the bottom half of Table 1.

Real object hindsight. The materials, procedure, and scoring were similar to those in Experiment 1 with the following exceptions. We used a green and white airplane and yellow sunglasses for test trials and a red car for practice. Participants completed only the Real object (pixel) hindsight task. In Experiment 1, each transparency contained a unique set of black dots and each unique set of dots covered 10% of the surface area of the sheet. In Experiment 2, each unique set of dots covered 5% of the surface area of the sheet. We made this alteration to facilitate baseline identification. This manipulation worked as intended: mean baseline identification in Experiment 1 was 8.23 (SEM = .101), and mean baseline identification in Experiment 2 was 6.54 (SEM = .131). Finally, the experimenter also reminded children not to peek when he hid the toys during the baseline trials and later asked participants to confirm that Ernie was not peeking when the experimenter hid the toys during the Hindsight trials.

Discrete hindsight. We developed six new questions that tapped a variety of educationally relevant domains. These included Astronomy (Planets), Biology (Fish), Botany (Herbs), Geometry (Shapes), Geography (Countries), and Language (Turkish). For each question, we asked children a Baseline question followed later in the testing session by a Hindsight question. Participants learned the correct answer to half these questions prior to answering the Hindsight question (experimental questions). For the remaining questions, participants simply tried to recall their original answer when asked the Hindsight question (control questions). This procedure controls for regression to the mean (see Pohl, 2004). Baseline questions assumed the form, ''which of these shapes do you think is a rhombus?'' Hindsight questions assumed the form, ''when I first showed you all of these shapes, which one did you think was a rhombus?'' In all but one of these Discrete Hindsight questions, we presented the correct answer among 11 distractors. For the remaining question (Planet), we presented the correct answer among eight distractors. In all cases, response options appeared one beside another in black ink on a laminated white sheet of 8.5×14 inch paper. Table 5 lists each of the Discrete Hindsight questions, the correct answers to these questions, the locations of the correct answers, and the number of distractors in these questions.

The reason we call these questions ''Discrete'' is because the correct answer and distractors are all qualitatively different (e.g., rhombus is different from triangle, circle, diamond, hexagon, etc.). Discrete Hindsight questions were scored as 1 if the child altered his or her Hindsight response to the correct answer and as 0 if the child did not. For example, if the child originally thought that the hexagon was a rhombus, and later claimed that he first thought that the rhombus was the rhombus, he would receive a score of 1. If the child pointed to his original response (or any response other than the rhombus), he received

Table 5
Question Type, Correct Answer, Location of Correct Answer, and Number of Distractors Used in the Discrete Hindsight Task in Experiment 2

Question Type	Correct Answer	Location of Correct Answer	Number of Distractors
Planets	Uranus	7	8
Fish	Razorfish	4	11
Herbs	Tarragon	3	11
Shapes	Rhombus	7	11
Countries	Spain	9	11
Language	Book [Turkish word is, ''kitap'']	5	11

Note. For each question type, participants answered a baseline question when naïve (e.g., ''which of these planets do you think is Uranus?'') and later a hindsight question (e.g., ''when I first showed you all of these planets, which one did you think was Uranus?'').

a score of 0. An overall Discrete Hindsight Bias score was calculated as the difference between the average hindsight bias score for the experimental questions and that for the control questions. Thus, higher scores indicated greater Discrete Hindsight Bias.

Continuous hindsight. Unlike the Discrete Hindsight questions, for which the answers and distractors were discrete items, we developed six new questions for which the answers and distractors were ordered on a continuous scale. The Continuous Hindsight questions included ones about size of body parts (Hand and Foot), reaching ability (Reaching), mental rotation plus size matching (Bar), direct size matching (Deer), and color matching (Color). We describe three of these questions here. The remaining questions can be obtained from the authors.

For the *Hand* question, children saw a laminated sheet containing nine handprints that increased in size from 68 mm × 58 mm (height × width) to 99 mm × 86 mm. The experimenter asked, ''which of these hands do you think fits your hand?'' For the *Reaching* question, children saw a laminated height chart on the wall containing 37 nondescript black lines spaced 16 mm apart. The experimenter asked, ''how high do you think you can reach on your tippy toes?'' Starting at the top line of the chart, the experimenter used a pointer and asked children whether they could reach that high. The experimenter then moved the pointer down three lines (48 mm) and repeated the question until the child indicated the height that she or he could reach. For the *Bar* question, children saw a laminated sheet containing 12 vertical bars of equal width (12 mm) that increased in height (from 77 mm to 95 mm). The experimenter held a bar (82 mm × 12 mm) horizontally and approximately one foot from the vertical bars, and asked, ''which of these bars do you think is the same size as the bar in my hand?'' In all cases, the experimenter encouraged children to answer these questions without actively solving them (e.g., they could not place their hand atop the handprints).

Later in the session, children learned the answers to these questions and answered the hindsight questions. For half the trials, children learned the answer *prior* to answering the hindsight question (experimental questions, e.g., child learns that her hand matches hand #4 and then is asked, ''when I first showed you these hands, which one did you think fit your hand?''). For the remaining trials, children learned the answer *after* answering the hindsight question (control questions).

We call these questions, ''Continuous'' because the correct answer and distractors are all related on a continuous scale. Continuous Hindsight questions were scored as follows. First, we calculated a bias score for each participant on each question using Pohl's (2004) equation: abs(Baseline − Correct Answer) − abs(Hindsight − Correct Answer). Next, for each question, we calculated a mean and standard deviation around these bias scores over all participants. Next, for each question, we calculated a z-score using the equation: (Bias Score − Mean of Bias Score) / SD of Bias Score. Finally, we calculated the mean of the z-scores across the three experimental questions and subtracted the mean across the three control questions. Thus, higher scores indicated greater Continuous Hindsight Bias.

Theory of mind. We used the same battery of ToM tasks from Experiment 1.

Inhibitory control. We retained the Day/Night task from Experiment 1. We replaced Bear/Dragon with the Dimensional Change Card Sort (Card Sort hereafter; Frye et al., 1995) to generalize our findings from Experiment 1. In the *Card Sort*, participants completed four practice trials in which they sorted by shape, followed by four training trials in which they sorted by color. In training trials, there was no conflict between the correct answer when sorting by shape and the correct answer when sorting by color. Participants then completed six test trials in which they sorted by color (preswitch test trials), followed by six test trials in which they sorted by shape (postswitch trials). In test trials, there was conflict between the correct answer when sorting by shape and the correct answer when sorting by color. Thus, in the postswitch trials, participants had to inhibit the sorting rule that they had used when sorting by color in the preswitch test trials. Participants received corrective feedback during the training trials only. Participants received a score from 0 to 6 representing the number of trials answered correctly on the last six test trials. We then calculated proportion correct out of six.

Working memory. Children completed the *Count and Label* (Gordon & Olson, 1998) task in which the experimenter presented participants with three objects (a toy snake, a brush, and a toy frog) and then proceeded to first count them while pointing to each (e.g., ''one, two, three''). Next the experimenter labeled each while pointing (e.g., ''snake, brush, frog''). Finally, the experimenter both counted and labeled the objects while pointing to each (e.g., ''One is a snake, two is a brush, three is a frog''). Next, the experimenter introduced three new objects (block, key, boat) and asked the child to do as he had done: first count, then label, then count and label the objects. Children performed this activity twice. We scored only the final trial from each attempt (count and label): children were scored as incorrect if they labeled all the objects first and then counted them in

turn or vice versa, or (more commonly) if they said "one is a block, one is a key, one is a boat." We averaged scores across the two trials.

Children also completed the *Backward Digit Span* (Davis & Pratt, 1995). The experimenter introduced children to a puppet, Ernie, and said that whatever he (experimenter) says, Ernie says it backward. The experimenter demonstrated, saying "1, 2" and then made Ernie say "2, 1." The experimenter asked children to do like Ernie (using the same example). The experimenter then asked that they do more like that, explaining that whatever the experimenter says, the children should say it backward. He began with two digits and increased the number of digits until children missed three consecutive trials. We recorded the highest level of success obtained (2, 3, 4, or 5 digits). Children received a score of 1 when they failed at 2 digits.

Language ability. Children completed the PPVT – III.

Procedure

We tested children individually in one videotaped session (lasting approximately one hour). Parents observed the sessions from an adjacent room. Children completed one block of 12 trials on the PPVT-III and then completed the remaining tasks in the following order. Children completed the Baseline trials from the Discrete Hindsight questions and then the Continuous Hindsight questions, followed by the Baseline trials of the Real Object Hindsight task. Children then completed three more blocks of the PPVT-III before completing all ToM tasks. This was followed by the remaining PPVT-III trials–testing ceased as soon as the child failed eight or more trials in a block. Next, children completed the Hindsight trials of the Real Object Hindsight task. Then children completed the Executive Function tasks, followed by the Hindsight trials of the Discrete and Continuous Hindsight tasks. The same male experimenter tested all children.

Counterbalancing

There were a total of four counterbalancing orders with 18 participants in each. For all four, the Discrete and Continuous Hindsight Bias questions occurred in a fixed order. However, we counterbalanced these questions such that half the items occurred with feedback (Hindsight condition) or without feedback (Control condition) prior to children answering the Hindsight questions. We also administered the ToM tasks in a fixed order. However, we counterbalanced these tasks for the Self and Other question: children either received the Self question first in the Unexpected Contents, Appearance/Reality, and Occluded Pictures, or they received the Other question first. We counterbalanced the Executive Function tasks such that half the children completed the Working Memory tasks first (Count and Label and then Backward Digit Span, in that order) prior to completing the Inhibitory Control tasks (Card Sort and then Day/Night, in that order) or vice versa.

Results and Discussion

As in Experiment 1, we address each research question in turn. Note that for all tasks except Day/Night, Count and Label, and Backward Digit Span ($N = 71$ for each) we had complete data for all 72 participants. In cases of missing data, we used list-wise deletion.

Preliminary Analyses

Proportion correct on each of the ToM, inhibitory control, and working memory tasks by age group appear in the bottom half of Table 2. Means and standard errors of each of the hindsight tasks by age group appear in Figure 3. When we combined variables measured on different scales, we averaged *z*-transformed values. As can be seen in Table 6, the ToM measures were interrelated, and therefore combined to form a ToM scale (Cronbach's $\alpha = .85$). The inhibitory control tasks were unrelated ($r = .12$) and therefore kept separate. The working memory tasks were related ($r = .57$) and therefore combined. For the Hindsight tasks, the Discrete and Continuous Hindsight were related ($r = .44$) and therefore combined (Discrete and Continuous Hindsight Bias hereafter). The Real Object Hindsight task was weakly related to both the Discrete Hindsight task ($r = .23$, $p = .05$) and the Continuous Hindsight task ($r = .16$, $p = .17$) and kept separate due to these weak correlations. Note that the data pattern remained when we combined all three Hindsight tasks to form a single scale. As can be seen in the bottom half of Table 4, the significant correlations between age and ToM and between age and the Inhibitory Control and Working Memory measures reflect age-related changes similar to those reported in the literature (e.g., Carlson et al., 2002).

Hindsight Bias

Three- and four-year-olds, but not five-year-olds demonstrated significant Real Object Hindsight Bias (see Figure 3) (Mean bias ± 95% Confidence Interval for 3-, 4-, and 5-year-olds, respectively. Note that mean bias values ± the 95% Confidence Interval that fall above 0 are significant: .24 ± .14; .30 ± .15; .07 ± .08). Unlike Experiment 1, the magnitude of the bias declined

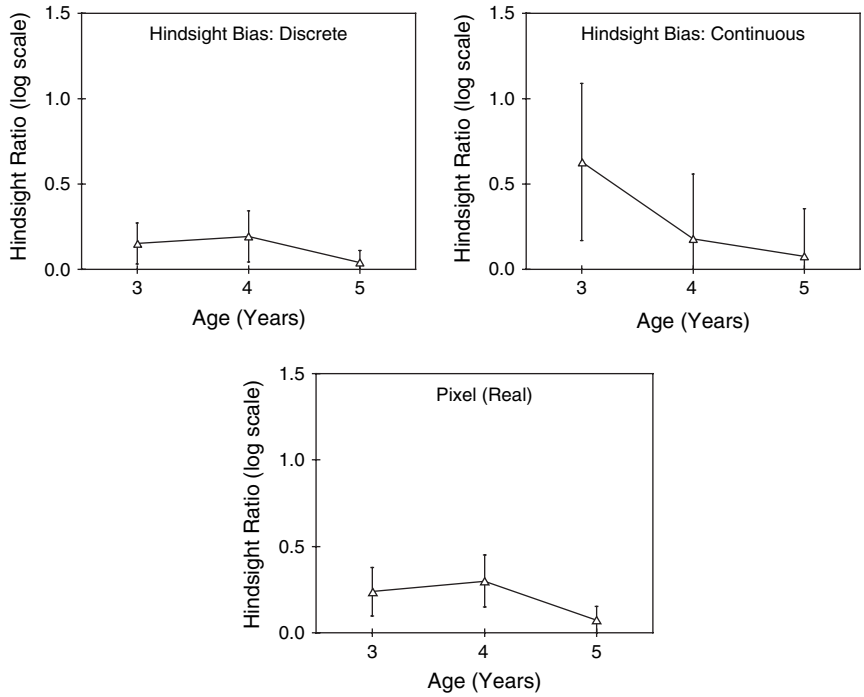

Figure 3. Magnitude of hindsight bias as a function of age in each of the three hindsight tasks in Experiment 2 (Discrete, Continuous, Pixel Real Object). Values greater than 0 indicate hindsight bias. Error bars indicate 95% confidence intervals.

significantly as a function of age $F(2,69) = 3.90$, $p = .025$. Post-hoc comparisons revealed that 3- and 4-year-olds showed more bias than 5-year-olds: $t(46) = 2.22$, $p = .031$ and $t(46) = 2.83$, $p = .007$, respectively. The two younger groups did not differ ($t < 1.0$).

Three- and four-year-olds, but not five-year-olds also demonstrated significant Discrete Hindsight Bias (Mean bias ± 95% Confidence Interval for 3-, 4-, and 5-year-olds, respectively. Note that mean Discrete and Continuous Hindsight bias values ± the 95% Confidence Interval that fall above 0 are significant: .15 ± .12; .19 ± .15; and .04 ± .07). The magnitude of this bias did not decline significantly with age $F(2,69) = 1.86$, $p = .16$. Only 3-year-olds demonstrated significant Continuous Hindsight Bias (Mean bias ± 95% Confidence Interval for 3-, 4-, and 5-year-olds, respectively: .63 ± .46; .18 ± .38; and .08 ± .28). The Continuous Hindsight Bias declined with age but did not reach conventional levels of significance $F(2,68) = 2.55$, $p = .08$.

Presently, we do not know why the Pixel Real Object Hindsight task yielded different data patterns in Experiments 1 and 2 (declining as a function of age in this experiment). One possibility is that we refined our procedure in Experiment 2; another is that the difference is due to random error. Replication of this task is, therefore, warranted.

Relation Between ToM and Hindsight Bias

As in Experiment 1, ToM correlated significantly with Hindsight Bias: ToM and Real Object Hindsight Bias ($r = -.36$, $p = .003$); ToM and Discrete and Continuous Hindsight Bias ($r = -.39$, $p = .001$, see bottom half of Table 4). To test more conservatively the relation between hindsight bias and ToM, we conducted a hierarchical multiple regression by entering age, then language ability, then executive function (Card Sort, Day/Night, and Working memory), and finally Hindsight Bias (Discrete and Continuous Hindsight Bias, Real Object Hindsight Bias) to predict ToM performance. Age (R^2 Change = .39, $p < .001$) and language (R^2 Change = .08, $p = .002$) accounted for significant variance, whereas executive function (R^2 Change = .05, $p = .067$) accounted for marginally significant variance in ToM performance after controlling for age and language ability. Importantly, as in Experiment 1, Hindsight Bias accounted for an additional 12% of the variance in ToM when entered last into the regression equation (R^2 Change = .12, $p < .001$. Note that the Discrete and Continuous Hindsight Bias and the Real Object Hindsight Bias each accounted for roughly 6% of the variance in ToM). All predictor variables together accounted for 65% of the variance in ToM performance. We reran this

Table 6
Zero-Order Correlations for the ToM, Inhibitory Control, Working Memory and Hindsight Tasks in Experiment 2

Task	CL	OP	AR	D/N	CS	CLA	BDS	PRO HB	DHB	CHB
UC	.52*	.79*	.47*	.20	.47*	.53*	.65*	−.32*	−.37*	−.21^
CL		.52*	.20	.22^	.07	.03	.33*	−.20^	−.40*	−.01
OP			.54*	.08	.35*	.30*	.56*	−.34*	−.49*	−.33*
AR				.11	.14	.45*	.48*	−.33*	−.21^	−.22^
D/N					.12	.16	.22*	−.31*	−.18	−.06
CS						.51*	.44*	−.01	−.13	−.20^
CLA							.57*	−.03	.01	−.19
BDS								−.20	−.31*	−.17
PRO HB									.23^	.16
DHB										.44*

Note. UC = Unexpected Contents; CL = Change in Location; OP = Occluded Pictures; AR = Appearance Reality; D/N = Day/Night; CS = Card Sort; CLA = Count and Label; BDS = Backward Digit Span; PRO HB = Pixel Real Object Hindsight Bias; DHB = Discrete Hindsight Bias. Critical questions for self and other were averaged for Unexpected Contents, Occluded Pictures, and Appearance Reality.
*$p < .05$.
^$p < .10$.

regression on each age group separately, and the data pattern remained. That is, hindsight bias accounted for 10% to 31% unique variance in ToM after controlling for language ability, Inhibitory Control, and Working Memory. Finally, we ran a regression equation in which we entered all predictor variables in one step. The resulting model had four predictor variables accounting for significant variance in ToM: Working Memory (*Beta* = .40, $p = .001$), age (*Beta* = .28, $p = .013$), Discrete and Continuous Hindsight Bias (*Beta* = −.24, $p = .004$), and Real Object Hindsight Bias (*Beta* = −.24, $p = .005$). In sum, hindsight bias accounted for significant variance in ToM performance, even after controlling for age, language ability, inhibitory control, and working memory.

Relation Between Inhibitory Control, Working Memory, and Hindsight Bias

Although Inhibitory Control, Working Memory, and Hindsight Bias all related to ToM, the three former constructs were largely unrelated to each other (see bottom half of Table 4). Card Sort and Working Memory correlated with ToM ($r = .41, p < .001$ and .67, $p < .001$, respectively), replicating previous work (Carlson & Moses, 2001; Carlson et al., 2002). Day/Night was unrelated to ToM, however ($r = .17$, $p = .14$). Based on the results of Experiment 1, we did not expect the two Inhibitory Control measures to correlate with Hindsight Bias. However, Day/Night correlated significantly with Real Object Hindsight Bias ($r = -.31$, $p = .008$), but it was not significantly related to Discrete and Continuous Hindsight Bias ($r = -.16$, $p = .19$). There were no other significant correlations.

The chief findings replicate and extend the findings of Experiment 1. Specifically, the results of Experiment 2 show that (a) preschoolers exhibit robust hindsight bias by claiming that a naïve peer or a naïve "prior" self knows more than is appropriate, (b) hindsight correlates significantly with performance on classic ToM tasks, (c) hindsight bias does not correlate with inhibitory control and working memory.

General Discussion

We developed a new battery of hindsight tasks appropriate for 3- to 5-year-old children and found robust hindsight bias. These tasks retained the basic structure of those administered to adults in cognitive and social psychology experiments (including both the hypothetical and memory design). There were significant correlations between children's performance on these hindsight tasks and their performance on classic ToM tasks. In two experiments, 3-, 4-, and 5-year-old children completed a battery of ToM tasks and measures of language ability, inhibitory control, working memory (Experiment 2 only), and hindsight bias. Results demonstrated that, as expected, the greater one's hindsight bias, the worse one's ToM performance. It is worth noting that our ToM tasks correlated with each other well (Cronbach's $\alpha = .85$ in both experiments), replicating prior work (e.g., Carlson & Moses, 2001; Wellman et al., 2001). Given this level of internal consistency, the *most* variance one could hope to account for in ToM performance

with any particular set of predictor variables is the Cronbach's α value squared ($.85^2$ = 72%). Using a combination of predictor variables, we accounted for 51% of the variance in ToM in Experiment 1 and 65% of the variance in Experiment 2. Thus, taken together, age, language ability, executive function (i.e., inhibitory control and working memory), and hindsight bias accounted for most of the variance that was statistically available to explain in ToM. The correlation between hindsight bias and ToM remained after we controlled for age, language ability, inhibitory control, and working memory. Perhaps most strikingly, hindsight bias accounted for numerically *more* variance in ToM performance than language ability and inhibitory control, two measures that have been found to account for significant variance in ToM performance (Astington & Jenkins, 1999; Carlson & Moses, 2001).

Contrary to what some have proposed (e.g., Birch & Bloom, 2004; Royzman et al., 2003), inhibitory control did not mediate the relation between hindsight bias and ToM. This finding should, however, be viewed cautiously given that the inhibitory control measures in both Experiments 1 (Day/Night and Bear/Dragon) and 2 (Day/Night and Card Sort) were not significantly intercorrelated. This may have reduced our ability to detect the contribution of inhibitory control to both ToM and hindsight bias.

As outlined in the Introduction, there are two related aspects to passing false belief tasks: (a) understanding that the mind can misrepresent reality, and (b) recapturing one's previous judgment in the face of new knowledge. Whereas ToM tasks involve both aspects, hindsight tasks involve only the second. From this perspective, it makes sense that ToM and hindsight bias show different developmental trajectories: children show substantial ToM gains between 3 and 5 years of age (in both the present and much previous work), while children do not show similarly dramatic developmental changes in hindsight bias (in Experiment 1 of the present work and in Bernstein et al., 2004). Although there was a slight decline in hindsight bias observed in Experiment 2 of the present work, it is clear that 4- and 5-year-olds do not "grow out" of hindsight bias in the same way they "grow out" of making errors on classic false belief tasks. In fact, we know from the adult literature that hindsight bias is a lifelong cognitive bias. In what follows, we first discuss possible reasons why ToM and hindsight tasks do not follow the same developmental trajectory and second, why, despite this, they are significantly related.

We find it helpful to think of our hindsight tasks as "False Hunch" or "False Judgment" tasks rather than classic "False Belief" tasks. In hindsight tasks, participants are asked their beliefs regarding matters about which they have little certainty. Participants state a belief about the situation in question, which they later must recall in the face of updated information. For example, in the adult case, participants are asked the height of the Statue of Liberty. Most adults do not know the answer, but guess (adults do not exhibit hindsight bias about facts for which they are certain). In the child case, participants are asked to state which shape is a representation of the country "Spain" or which color they think is "cyan" or what lies behind a pixilated pattern of visual noise. Again, the answers they provide to these questions are not held with certainty (see Pohl, 2004). Thus, an individual's initial knowledge state in a hindsight task is more like a hunch or a guess, rather than a well-grounded belief. Stated another way, in hindsight tasks the participant's first report is a judgment with relatively low subjective confidence.

In contrast, in classic false belief tasks, the participant is asked to recall a prior belief that was held with great certainty. In the unexpected contents false belief task, the child sees a crayon box. The participant strongly and reasonably believes that the box contains crayons—this is not a wild guess. Even more clearly in the change in location false belief task, the child sees an object put in one location and later, after it is moved in the absence of Billy, the child is asked where Billy will look for it. Before the child is given information about object movement, she has full knowledge that it is in Location 1 and holds this belief with great certainty.

Because young children do poorly on both hindsight and false belief tasks, and older children and adults master false belief tasks yet continue to do poorly on hindsight tasks, the certainty with which one holds an initial belief state (i.e., how "committed" one is to this belief) seems to be a relevant factor to consider when trying to explain differences in performance on these two tasks.

One reason why hindsight bias may persist throughout life is that information that we hold with low certainty is highly malleable and profoundly influenced by new information about reality. This "mental tagging" of uncertainty about information allows for facile updating (which is a handy feature of cognition; see Gigerenzer & Goldstein, 1996) but also allows us to fall prey to hindsight bias (the updating prevents gaining full and easy access to our past beliefs). In short, the fact that we hold an idea with a "low certainty tag" may allow or invite easy updating based on feedback, which, in turn, may contribute both to avid learning (which is good) and also the "I knew it all along" effect (which is a consequent

cognitive 'bug'). With respect to ToM, the fact that we hold an idea with a ''high certainty tag'' may prevent easy updating based on feedback, thereby reducing interference and helping adults, at least, to gain full access to past beliefs. This combination of factors may be why older children and adults can access strongly held past beliefs that are contradicted by the current reality (success on false belief tasks) while still doing poorly on hindsight tasks.

Despite the fact that performance on the ToM and hindsight tasks did not show similar developmental patterns, our data clearly indicate a modest relation between ToM and hindsight bias. Given the correlational nature of this study, we cannot make causal claims (e.g., whether holding a more nuanced ToM contributes to a reduction in hindsight bias, or the reverse). However, the negative correlation between ToM and hindsight bias is in the direction that one would expect: If a child did not understand that his/her own beliefs could be false or unduly influenced by current reality, one would expect the child to do poorly on hindsight tasks—the child would naturally think he ''knew it all along.'' Admittedly, this falls short of a satisfactory and full account of the reasons for the observed correlations between ToM and hindsight bias performance, but the negative direction of the correlation is understandable.

On a more speculative level, the notion of ''fluency'' introduced by cognitive scientists may provide another perspective on the link between hindsight bias and ToM. Prior exposure to stimuli (words, pictures, sounds) improves the fluency (i.e., speed, effort, and accuracy) with which people later process those stimuli (Jacoby, Kelley, & Dywan, 1989). People can experience fluency as familiarity, and this can affect many judgments, including memory, preference, and truth (see Bernstein, 2005; Bernstein, Whittlesea, & Loftus, 2002; Brinol, Petty, & Tormala, 2006; Clore, 1992). Harley and colleagues suggested a ''processing fluency account'' of hindsight bias in adults (Bernstein & Harley, 2007; Harley et al., 2004; see also Roese, Fessel, Summerville, Kruger, & Dilich, 2006; Sanna, Schwarz, & Small, 2002). According to this account, the outcome to a problem is salient and thus processed fluently. This fluent knowledge, in turn, is hard to ignore when estimating what a naïve self or other knows. For example, in the current work, participants watched common objects clarify on a computer screen or behind a series of filter screens in a Baseline condition. Later, in a Hindsight condition, participants had to estimate when a naïve peer would identify those same objects as they clarified. Once known, the objects are processed fluently. As Jacoby (1978) demonstrated, sometimes it is easier to retrieve an answer to a problem than it is to re-solve the problem. In the hindsight task, participants must ignore or discount this fluency to correctly estimate when a naïve peer will identify the object. Of course, ''fluency'' more naturally accounts for hindsight bias than for ToM, but future work might usefully examine cognitive fluency, as described in the cognitive and social psychology literatures, as a link between the errors committed in hindsight bias and various ToM tasks.

Regardless of how exactly hindsight bias relates to ToM, we maintain that our new hindsight tasks provide useful tools for exploring perspective-taking errors across the lifespan. Our Computer Hindsight Task has already been employed in other tests of preschoolers and adults (Bernstein et al., 2004). Our Real Object Hindsight task correlates with the Computer Hindsight task, and both correlate with ToM tests. Finally, our two newest hindsight bias measures, the Discrete Hindsight Task and the Continuous Hindsight Task, also correlate with ToM. Thus, all of our hindsight tasks correlate with ToM. One advantage of our hindsight tasks is that they are continuous measures (Appleton-Knapp, 2002; Guilbault, Bryant, Brockway, & Posavac, 2004). Classic false belief tasks are discrete measures: One either passes or fails each task. Such discrete coding may limit one's ability to capture important individual variability. Continuous hindsight tasks may be useful in detecting subtle variations in perspective taking that go undetected in the standard battery of tests typically employed in ToM research.

Conclusion

In sum, our findings reveal a kinship between hindsight bias and ToM. Despite this kinship, these two constructs follow different developmental trajectories: Children master classic false belief tasks by age 5, whereas people, even adults, continue to exhibit hindsight bias throughout life. The time is ripe for further studies examining the relation between hindsight bias and ToM. Doing so would allow for important contributions to what Flavell, Miller, & Miller (2002) have referred to as the *horizontal* ties of a theory of mind. Whereas *vertical* ties refer to the connections between sequences of theory of mind developments (e.g., how early understanding of mind transforms into subsequent understanding), horizontal ties refer to the relation between theory of mind and other concepts developing concurrently (e.g., emotional development and other forms of perspective taking). By exploring both vertical and horizontal ties, we will enrich our understanding of the developmental relation between hindsight bias and ToM

and expand our understanding of human social cognition.

References

Appleton-Knapp, S. L. (2002). *Memory dynamics in hindsight bias. Unpublished doctoral dissertation*.

Arkes, H. R., Wortman, R. L., Saville, P. D., & Harkness, A. R. (1981). Hindsight bias among physicians weighing the likelihood of diagnoses. *Journal of Applied Psychology, 66*, 252–254.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology. 35*, 1311–1320.

Atance, C., & Meltzoff, A. N. (2006). Preschoolers' current desires warp their choices for the future. *Psychological Science*, *17*, 583–587.

Bernstein, D. M. (2005). Making sense of memory. *Canadian Journal of Experimental Psychology, 59*, 199–208.

Bernstein, D. M., Atance, C., Loftus, G. R., & Meltzoff, A. N. (2004). We saw it all along: Visual hindsight bias in children and adults. *Psychological Science*, *15*, 264–267.

Bernstein, D. M., & Harley, E. M. (2007). *Fluency misattribution and visual hindsight bias*. Manuscript in press.

Bernstein, D. M., Loftus, G. R., & Meltzoff, A. N. (2005). Object identification in preschool children and adults. *Developmental Science*, *8*, 151–161.

Bernstein, D. M., Whittlesea, B. W. A., & Loftus, E. F. (2002). Increasing confidence in remote autobiographical memory and general knowledge: Extensions of the revelation effect. *Memory & Cognition*, *30*, 432–438.

Birch, S. A. J. (2005). When knowledge is a curse: Children's and adults' reasoning about mental states. *Current Directions in Psychological Science*, *14*, 25–29.

Birch, S. A. J., & Bernstein, D. M. (2007). What can tell us about hindsight bias: A fundamental constraint on perspective children taking? *Social Cognition 25*, 98–113.

Birch, S. A. J. & Bloom, P. (2003). Children are cursed: An asymmetric bias in mental-state attribution. *Psychological Science*, *14*, 283–286.

Birch, S. A. J. & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Science*, *8*, 255–260.

Brinol, P., Petty, R. E., & Tormala, Z. L. (2006). The malleable meaning of subjective ease. *Psychological Science*, *17*, 200–206.

Camerer, C., Loewenstein, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy, 97*, 1232–1254.

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28*, 595–616.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*, 1032–1053.

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development*, *11*, 73–92.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development*, *69*, 672–691.

Clore, G. L. (1992). Cognitive phenomenology: Feelings and the construction of judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 133–163). Hillsdale, NJ: Lawrence Erlbaum.

Davis, H. L., & Pratt, C. (1995). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology, 47*, 25–31.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology, 40*, 760–768.

Fischhoff, B. (1975). Hindsight $\neq$ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 288–299.

Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology, 15*, 95–120.

Flavell, J. H., Green, F. L., & Flavell, E. R. (1986). Development of knowledge about the appearance-reality distinction. *Monographs of the Society for Research in Child Development*, *51* (1, serial no. 212).

Flavell, J. H., Miller, P. H., & Miller, S. A. (2002). *Cognitive development*. Upper Saddle River, NJ: Prentice Hall.

Friedman, O., & Leslie, A. M. (2004). Mechanisms of belief-desire reasoning: Inhibition and bias. *Psychological Science*, *15*, 547–552.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, *10*, 483–527.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*, 650–669.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance reality distinction. *Child Development*, *59*, 26–37.

Gordon, A. C. L., & Olson, D. R. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology, 68*, 70–83.

Guilbault, R. L., Bryant, F. B., Brockway, J. H., & Posavac, E. J. (2004). A meta-analysis of research on hindsight bias. *Basic and Applied Social Psychology, 26*, 103–117.

Harley, E. M. (2007). Hindsight bias in legal decision making. *Social Cognition*, *25*, 48–63.

Harley, E. M., Carlsen, K., & Loftus, G. R. (2004). The ''saw it all along effect'': Demonstrations of visual hindsight bias. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*, 960–968.

Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, *107*, 311–327.

Hoffrage, U., Hertwig, R., & Gigerenzer, G. (2000) Hindsight bias: A by-product of knowledge-updating?, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 566–581.

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

Jacobs, J. E., & Klaczynski, P. A. (2002). The development of judgment and decision making during childhood and adolescence. *Current Directions in Psychological Science*, *11*, 145–149.

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior*, *17*, 649–667

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger & F. I. M. Craik (Eds.). *Varieties of memory and consciousness: Essays in honor of Endel Tulving* (pp. 391–422). Hillsdale, NJ: Erlbaum.

Kamin, K. A., & Rachlinski, J. J. (1995). Ex post $\neq$ ex ante: Determining liability in hindsight. *Law and Human Behavior*, *19*, 89–104.

Keenan, T. (1998). Memory span as a predictor of false belief understanding. *New Zealand Journal of Psychology*, *27*, 36–43.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*, 25–41.

Leary, M. R. (1981). The distorted nature of hindsight. *Journal of Social Psychology*, *115*, 25–29.

Leary, M. R. (1982). Hindsight distortion and the 1980 Presidential election. *Personality and Social Psychology Bulletin*, *8*, 257–263.

Leslie, A. M., Friedman, O., & German, T. (2004). Core mechanisms in 'theory of mind.' *Trends in Cognitive Science*, *12*, 528–533.

Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology*, *50*, 45–85.

Mitchell, P., Robinson, E. J., Isaacs, J. E., & Nye, R. M. (1996). Contamination in reasoning about false belief: An instance of realist bias in adults but not children. *Cognition*, *59*, 1–21.

Moses, L. J. (2001). Executive accounts of theory-of-mind development. *Child Development*, *72*, 688–690.

Nickerson, R. S. (1999). How we know–and sometimes misjudge–what others know: Imputing one's own knowledge to others. *Psychological Bulletin*, *126*, 747–753.

Perner, J., & Lang, B. (2000). Theory of mind and executive function: Is there a developmental relationship? In S. Baron-Cohen, H. Tager-Flúsberg, D. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (2nd ed., pp. 150–181, UK). Oxford, UK: Oxford Unversity Press.

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, *5*, 125–137.

Pohl, R. F. (2004). Hindsight bias. In R. F. Pohl (Ed.), *Cognitive illusions*: *A handbook on fallacies and biases in thinking, judgement and memory* (pp. 363–368). Hove, UK: Psychology Press.

Pohl, R. F., Eisenhauer, M., & Hardt, O. (2003). SARA: A cognitive process model to simulate the anchoring effect and hindsight bias. *Memory*, *11*, 337–356.

Pohl, R. F., & Haracic, I. (2005). Der Rückschaufehler bei Kindern und Erwachsenen [Translated Title: Hindsight bias in children and adults]. *Zeitschrift fur Entwicklungspsychologie und Padagogische Psychologie*, *37*, 46–55.

Pohl, R. F. & Hell, W. (1996). No reduction in hindsight bias after complete information and repeated testing. *Organizational Behavior & Human Decision Processes*, *67*, 49–58.

Roese, N. J., Fessel, F., Summerville, A., Kruger, J., & Dilich, M. A. (2006). The propensity effect: When foresight trumps hindsight. *Psychological Science*, *17*, 305–310.

Royzman, E. B., Cassidy, K. W., & Baron, J. (2003). ''I know, you know'': Epistemic egocentrism in children and adults. *Review of General Psychology*, *7*, 38–65.

Russell, J. (1996). *Agency: Its role in mental development*. Hove, UK: Erlbaum.

Sanna, L. J., Schwarz, N., & Small, E. M. (2002). Accessibility experiences and the hindsight bias: I knew it all along versus it could never have happened. *Memory & Cognition*, *30*, 1288–1296.

Scullin, M. H., & Bonner, K. (2006). Theory of mind, inhibitory control, and preschool-age children's suggestibility in different interviewing contexts. *Journal of Experimental Child Psychology*, *93*, 120–138.

Taylor, M. (1988). Conceptual perspective taking: Children's ability to distinguish what they know from what they see. *Child Development*, *59*, 703–718.

Taylor, M., Esbensen, B. M., & Bennett, R. T. (1994). Children's understanding of knowledge acquisition: The tendency for children to report that they have always known what they have just learned. *Child Development*, *65*, 1581–1604.

Taylor, M., & Flavell, J. H. (1984). Seeing and believing: Children's understanding of the distinction between appearance and reality. *Child Development*, *55*, 1710–1720.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*, 655–684.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*, 523–541.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong

beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128.

Wood, G. (1978). The knew-it-all-along effect. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 345–353.

Zelazo, P. D., Mueller, U., Frye, D., & Marcovitch, S. (2003). *The development of executive function in early childhood. Monographs of the Society for Research in Child Development, 68 (3, Serial No. 274).*

Zwick, R., Pieters, R., & Baumgartner, H. (1995). On the practical significance of hindsight bias: The case of the expectancy-disconfirmation model of consumer satisfaction. *Organizational Behavior and Human Decision Processes*, *64*, 103–117.