

# Journal of Experimental Psychology: Learning, Memory, and Cognition

## Hindsight Bias From 3 to 95 Years of Age

Daniel M. Bernstein, Edgar Erdfelder, Andrew N. Meltzoff, William Peria, and Geoffrey R. Loftus

Online First Publication, February 7, 2011. doi: 10.1037/a0021971

### CITATION

Bernstein, D. M., Erdfelder, E., Meltzoff, A. N., Peria, W., & Loftus, G. R. (2011, February 7). Hindsight Bias From 3 to 95 Years of Age. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. doi: 10.1037/a0021971

# Hindsight Bias From 3 to 95 Years of Age

Daniel M. Bernstein

Kwantlen Polytechnic University and University of Washington

Edgar Erdfelder

University of Mannheim

Andrew N. Meltzoff, William Peria, and Geoffrey R. Loftus

University of Washington

Upon learning the outcome to a problem, people tend to believe that they knew it all along (*hindsight bias*). Here, we report the first study to trace the development of hindsight bias across the life span. One hundred ninety-four participants aged 3 to 95 years completed 3 tasks designed to measure visual and verbal hindsight bias. All age groups demonstrated hindsight bias on all 3 tasks; however, preschoolers and older adults exhibited more bias than older children and younger adults. Multinomial processing tree analyses of these data revealed that preschoolers' enhanced hindsight bias resulted from them substituting the correct answer for their original answer in their recall (a qualitative error). Conversely, older adults' enhanced hindsight bias resulted from them forgetting their original answer and recalling an answer closer to, but not equal to, the correct answer (a quantitative error). We discuss these findings in relation to mechanisms of memory, perspective taking, theory of mind, and executive function.

*Keywords:* hindsight bias, life span cognitive development, mathematical models of cognition, executive function, perspective taking

At 8:46 a.m. and 9:03 a.m. on September 11, 2001, American Airlines Flight 11 and United Airlines Flight 175, respectively, crashed into the north and south towers of the World Trade Center in New York City, New York, killing everyone on board and an unknown number of people in the towers (National Commission on Terrorist Acts Upon the United States, 2004 [hereafter, the 9/11 Commission]). The 9/11 Commission was given the unenviable task of determining what was known prior to September 11, 2001, and whether this foreknowledge could have been used to prevent the terrorist attacks of that day. To their credit, the Commission

authors were well aware of how difficult their task was. They titled Chapter 11 of their report "Foresight—And Hindsight" and began that chapter as follows:

In composing this narrative, we have tried to remember that we write with the benefit and the handicap of hindsight. Hindsight can sometimes see the past clearly—with 20/20 vision. But the path of what happened is so brightly lit that it places everything else more deeply into shadow. Commenting on Pearl Harbor, Roberta Wohlstetter found it "much easier *after* the event to sort the relevant from the irrelevant signals. After the event, of course, a signal is always crystal clear; we can now see what disaster it was signaling since the disaster has occurred. But before the event it is obscure and pregnant with conflicting meanings." (9/11 Commission, 2004, p. 339)

---

Daniel M. Bernstein, Department of Psychology, Kwantlen Polytechnic University, Surrey, British Columbia, Canada, and Institute for Learning & Brain Sciences, University of Washington; Edgar Erdfelder, Department of Psychology, University of Mannheim, Mannheim, Germany; Andrew N. Meltzoff, Institute for Learning & Brain Sciences, University of Washington; William Peria and Geoffrey R. Loftus, Department of Psychology, University of Washington.

This research was supported by National Science Foundation Science of Learning Center Grant SBE-0354453, National Institutes of Health Grant MH41637, and Social Sciences and Humanities Research Council of Canada Grant 410-2008-1681. The views expressed here are the authors' and are not necessarily those of the funding agencies. We thank Joy Durham for help with data collection and thank the University Child Development School, Langley Gardens, Arborside Court, and the 411 Seniors Centre Society for allowing us to test at their sites. Portions of this work were presented at the meetings of the Psychonomic Society in 2008 and the Society for Research in Child Development in 2009.

Correspondence concerning this article should be addressed to Daniel M. Bernstein, Department of Psychology, Kwantlen Polytechnic University, 12666-72nd Avenue, Surrey, British Columbia, V3W 2M8 Canada. E-mail: daniel.bernstein@kwantlen.ca

Five years after the tragedy of 9/11, another tragedy, this one the result of a natural disaster, befell the United States. The day after Hurricane Katrina struck New Orleans, Louisiana, devastating the city, Michael Brown, former Federal Emergency Management Agency director, recalled his feelings from the day before Katrina hit: "I knew in my gut this was the bad one" (MSNBC News, 2006)—thus exemplifying the finding that, with the benefit of outcome knowledge, people claim that they or another person knew it all along (Fischhoff, 1975; Wood, 1978). Director Brown assumed, for example, that he could have predicted Hurricane Katrina's devastation. This *hindsight bias* lends the world's events an air of predictability and inevitability.

Hindsight bias is a common and robust cognitive error that has been documented across cultures (Pohl, Bender, & Lachmann, 2002) in a variety of judgments, including medical diagnoses (Arkes, Wortman, Saville, & Harkness, 1981), legal decisions (Harley, 2007), consumer satisfaction (Zwick, Pieters, & Baumgartner, 1995), sporting events (Leary, 1981), and election outcomes (Blank, Fischer, & Erdfelder, 2003; Leary, 1982). In each

case, people who have knowledge of an outcome overestimate the likelihood of that outcome. Hindsight bias provides the illusion of understanding the past and can result in a failure to learn from the past (Fischhoff, 1982).

What psychological mechanisms underlie hindsight bias? Earlier accounts posited that people automatically update their knowledge with new information, rendering the original information inaccessible (Fischhoff, 1975; see Blank & Nestler, 2007). Newer theories posit 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 (Dehn & Erdfelder, 1998; Hawkins & Hastie, 1990; see also Blank, Nestler, von Collani, & Fischer, 2008, for the view that there are at least three different components of hindsight bias, each subsumed by different processes).

Hindsight bias studies often involve almanac questions (e.g., “How many books did Agatha Christie write?”) given to adults. Few studies have examined hindsight bias in preschool children and older adults. One reason for the lack of child developmental work on hindsight bias is due to numerical and linguistic limitations in preschoolers. Despite these limitations, researchers have developed other ways to measure hindsight bias in young children. These include measuring a knowledgeable child’s estimates of what a naïve peer would know about the contents of a toy or about the visual identity of degraded objects (Bernstein, Atance, Loftus, & Meltzoff, 2004; Bernstein, Atance, Meltzoff, & Loftus, 2007; Birch & Bloom, 2003). From the few extant studies, there is mixed evidence for developmental changes in hindsight bias in the preschool years. There is some evidence for a developmental decline in hindsight bias from childhood to adulthood (Pohl, Bayen, & Martin, 2010), but evidence here too is mixed (see Bayen, Pohl, Erdfelder, & Auer, 2007; Birch & Bernstein, 2007, for reviews). The evidence for hindsight bias in older adults is also sparse. In the only published study of which we are aware, younger and older adults answered almanac questions and later learned the answers to half the questions prior to recalling their original answers. As expected, both age groups showed hindsight bias by recalling answers closer to the actual answers that had been provided; however, older adults showed more bias than younger adults (Bayen, Erdfelder, Bearden, & Lozito, 2006).

In sum, the limited evidence to date indicates that preschoolers and older adults may be more prone to hindsight bias than are older children and younger adults. As Bayen and colleagues (2007) noted,

A shortcoming of the existing literature is that empirical studies either compared children with younger adults, or older adults with younger adults, but thus far no common methodology to address developmental issues across the entire life span has been developed. (p. 84)

The current work redresses this shortcoming.

Few developmental studies have utilized identical measures to assess cognition in preschoolers, older children, adults, and older adults. We believe that developing and using the same measures on children and adults can inform theories of cognitive development (Meltzoff, Kuhl, Movellan, & Sejnowski, 2009). We acknowledge that performance on the same sets of measures in children and adults does not imply the same underlying cognitive process(es). However, we assert that to the extent that the same tasks can be

used for all age groups studied, researchers can identify cognitive processes of interest. Conversely, we feel that using different tasks when one is attempting to measure the same underlying construct(s) at different ages, for example, executive function, presents interpretational difficulties. We sought to avoid these difficulties here by creating a battery of three tasks that could be administered in the same form to participants across a wide age range—from 3 to 95 years of age.

There are several reasons to study a robust cognitive error like hindsight bias across the life span. First, life span developmental studies can inform theory and practice in terms of how and when biases form and change (Baltes, 1987). Second, hindsight bias may relate to research in developmental psychology that is generally called *theory of mind* (ToM; e.g., Perner, 1991; Wellman, 1990). ToM research is a broad area that investigates children’s understanding that other minds are different from one’s own mind; hence, oneself and others can hold mistaken or false beliefs about the world (Astington & Gopnik, 1991; Flavell, 1999). ToM and hindsight bias correlate modestly in preschoolers and young school-age children after controlling for variables known to relate to ToM, including age, language ability, and executive function (Bernstein et al., 2007). A third reason to study hindsight bias across the life span is that mathematical modeling of cognitive-developmental data, using an approach called multinomial processing trees (MPTs), can reveal the different components involved in hindsight bias and determine how they differ by age. Specifically, MPT modeling can reveal differences in cognitive development that arise from either qualitative or quantitative changes. In the context of hindsight bias, qualitative changes refer to the kind or quality of errors that children and adults make; in contrast, quantitative changes refer to the number or quantity of errors that children and adults make. By modeling hindsight bias data from a large life span developmental cross-sectional study, we might learn more about the nature of hindsight bias in particular and memory in general.

In the current work, we sought to (a) chart the development of hindsight bias across the life span and (b) reveal the mechanisms underlying hindsight bias. We administered a set of verbal and visual hindsight bias tasks to preschoolers, elementary and middle school children, younger adults, and older adults, and used MPT models to address these aims. Despite the difficulties discussed previously regarding verbal hindsight bias tasks in preschoolers, such tasks are among the most standard. We therefore used a verbal task, allowing us to connect our results to the broader hindsight bias literature. In the verbal hindsight task, participants answered general-knowledge questions. Later, participants either learned the correct answers to half the questions (experimental questions) or not (control questions) before being asked to recall their original answers. The extent to which participants recalled their original answers as being closer to the correct answers in the experimental condition than in the control condition reflects verbal hindsight bias (see Pohl, 2007).

Following Bernstein et al. (2007), we also administered two visual hindsight tasks, one involving computerized images of common objects (computer hindsight) and the other involving real objects hidden behind a series of filter screens (real object hindsight). In both tasks, during a foresight judgment (FJ) condition, participants tried to identify objects as the objects gradually clarified. Later, during a hindsight judgment (HJ) condition, partici-

pants learned those same objects' identities at the start of the clarification process and tried to estimate when a same-age naïve peer named Ernie would identify the objects. The extent to which participants identified objects at a clearer state in the FJ condition than in the HJ condition reflects visual hindsight bias.

Thus, our verbal hindsight task follows a memory design in which participants answer questions at Time 1 and then later try to recall their answers at Time 2 either in the presence or absence of learning the correct answers. Our visual hindsight task follows what we call a quasi-hypothetical design: Participants identify objects at Time 1 and then later at Time 2, in the presence of learning the identity to each of those objects, participants try to estimate when a naïve peer will identify the objects. The reason that it is quasi-hypothetical is that in the standard hypothetical design in hindsight bias experiments, there is no baseline set of trials at Time 1; rather, participants simply learn the correct answer or not and then estimate what a naïve peer would say the answer was. Both verbal hindsight bias and visual hindsight bias represent the extent to which outcome knowledge colors one's judgments about either one's former or another person's naïve knowledge state.

### Modeling Hindsight Bias Using Multinomial Processing Trees

Hindsight bias theories tend to focus on recollection or reconstruction biases. *Recollection bias* occurs when outcome knowledge either alters one's memory for his or her original judgment (OJ; Fischhoff, 1975; E. F. Loftus & Loftus, 1980) or renders one's OJ inaccessible (Hell, Gigerenzer, Gauggel, Mall, & Müller, 1988). *Reconstruction bias* occurs when one fails to recall one's OJ and uses the outcome knowledge to help reconstruct one's OJ.

Our goal here was to track age-related changes in different mechanisms believed to underlie hindsight bias, including recollection and reconstruction bias (Bayen et al., 2006; Pohl et al., 2010). To accomplish this, we used an MPT model approach to hindsight bias suggested by Erdfelder and Buchner (1998) and applied this approach to verbal and visual tasks. MPT models can be used to estimate probabilities of unobservable events from frequencies of observable events. MPT models have been used to provide pure measures of the cognitive processes assumed to underlie human judgments in various tasks (see Batchelder & Riefer, 1999; Erdfelder et al., 2009).

### Verbal Hindsight Bias as Measured by HB13

Erdfelder and Buchner (1998) presented an MPT model of hindsight bias that permits investigators to assess two fundamental ways in which people can be biased in hindsight. First, recollection bias arises when individuals fail to recall their original prediction after learning the outcome to an event. Second, reconstruction bias arises when, after learning the outcome to an event and failing to recall one's original predictions, outcome knowledge biases one's reconstruction of past predictions.

Using Erdfelder and Buchner's (1998) terminology, assume that Sally is asked how many keys there are on a piano. She answers 50. This is Sally's OJ. Later Sally learns the correct answer to this question (CJ, for correct judgment). In this case, CJ = 88 keys on a piano. At this point, after Sally has generated her OJ and has

learned the CJ, she is asked to recall her OJ or how many keys she originally thought were on a piano. We call this the recall of OJ (ROJ). Say that Sally's ROJ = 75. This means that Sally's recall of her original answer is still incorrect, but it is closer to the actual answer (OJ = 50; CJ = 88; ROJ = 75). This is a typical response in a hindsight bias study.

In addition to questions for which participants learn the CJ before trying ROJ, there are questions for which participants do not learn the CJ before attempting their ROJ. The former is the experimental condition, and the latter is the control condition. Overall hindsight bias is measured as the difference in bias scores between the experimental and control conditions, for example, using the bias measure  $\text{hindsight bias} = [(OJ - ROJ)/(OJ - CJ)] \times 100$  originally proposed by Hell et al. (1998).

To conduct an MPT analysis of hindsight bias data, we need to convert our continuous hindsight bias measure into a discrete measure. Erdfelder and Buchner (1998) proposed that researchers rank order their data corresponding to the three numerical quantities, OJ, CJ, and ROJ. When we do this, we see that there are five rank orders corresponding to the situation in which  $OJ < CJ$ —(a)  $ROJ < OJ < CJ$ , (b)  $ROJ = OJ < CJ$ , (c)  $OJ < ROJ < CJ$ , (d)  $OJ < ROJ = CJ$ , and (e)  $OJ < CJ < ROJ$ —and five rank orders corresponding to the situation in which  $CJ < OJ$ —(f)  $ROJ < CJ < OJ$ , (g)  $ROJ = CJ < OJ$ , (h)  $CJ < ROJ < OJ$ , (i)  $CJ < ROJ = OJ$ , and (j)  $CJ < OJ < ROJ$  (Bayen et al., 2006; Erdfelder & Buchner, 1998).

We can now use these 10 distinct OJ–CJ–ROJ rank orders to determine whether the ROJ deviates from the OJ toward the CJ (Rank Orders c, d, e, h, i, and j), away from CJ (Rank Orders a and f), or whether  $ROJ = OJ$  (Rank Orders b and g). We observe the frequencies of these 10 rank orders separately for control items and experimental items. These  $2 \times 10 = 20$  frequencies are the raw data to which Erdfelder and Buchner's (1998) hindsight bias model refers.

Figure 1 shows a simplified sketch of Erdfelder and Buchner's (1998, pp. 392–393) full MPT model of hindsight bias. The full model contains 13 parameters and is called the *HB13 model*. For our purposes, we focus on four core parameters of the model, and thus, Figure 1 suffices for illustration. These parameters are  $r_C$ ,  $r_E$ ,  $b$ , and  $c$ . Parameters  $r_C$  and  $r_E$  represent the probabilities of OJ recollection for control items (without feedback) and experimental items (with feedback), respectively. Parameter  $b$  reflects reconstruction bias (i.e., the probability of a CJ-biased reconstruction given recollection failure) and parameter  $c$  the probability of source confusion (i.e., reconstructing the CJ rather than the OJ for experimental items).

Using these four core parameters, we can determine the contributions of recollection and reconstruction biases to overall hindsight bias in each of our age groups. First, to determine whether recollection bias has occurred, we look for evidence that  $r_C > r_E$ . If this outcome occurs, then we conclude that recollection of experimental items has been hampered by knowledge of the CJs. Next, we can look for evidence of reconstruction bias. If reconstruction bias has occurred, then we will see that participants' ROJs are closer to the CJs for experimental compared to control items. Parameter  $b$  reflects this outcome. Last, we can assess the proportion of CJ-biased reconstructions in which the CJ is confused with the OJ. Parameter  $c$  represents the probability of such source confusions.

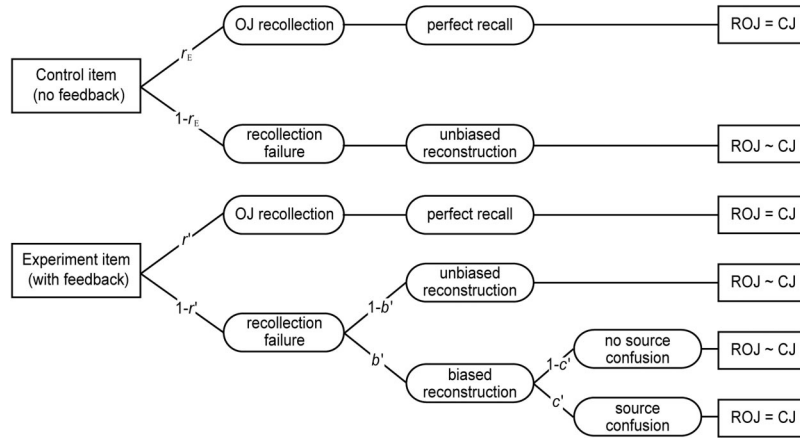


Figure 1. A sketch of the HB13 model for analyzing verbal hindsight bias data. Parameters  $r_C$  and  $r_E$  denote the probabilities of OJ recollection in the control condition and the experimental condition, respectively. Parameter  $b$  is the probability of biased OJ reconstruction given recollection failure. Parameter  $c$  is the probability of a source memory confusion in which  $ROJ = CJ$ . Rectangles signify observable events; ovals signify unobservable cognitive processes. Parameters denote transition probabilities between processes from left to right. OJ = original judgment; ROJ = recall of original judgment; CJ = correct judgment;  $\sim$  = is close to. Adapted from “Recollection Biases in Hindsight Judgments,” by E. Erdfelder, M. Brandt, and A. Bröder, 2007, *Social Cognition*, 25, p. 117. Copyright 2007 by Guilford Publications. For the full version of the HB13 model, see Erdfelder and Buchner (1998, p. 392).

To estimate all parameters of the HB13 model, we derive model equations by multiplying the probabilities that lead to a particular rank-order event. If more than one branch of a tree leads to the same rank-order event, the branch probabilities are summed. This way, we arrive at 20 model equations for the  $2 \times 10$  rank-order probabilities of experimental and control items, respectively (Bayen et al., 2006; Erdfelder & Buchner, 1998). On the basis of these model equations, parameters are estimated using the maximum-likelihood (ML) method. The likelihood ratio chi-square statistic  $G^2$  is used to test the model’s fit to a set of data (Hu & Batchelder, 1994).

### Visual Hindsight Bias as Measured by VHB3

In addition to examining the mechanisms underlying verbal hindsight bias using the HB13 model, we developed a visual hindsight bias model (VHB3). We tried to make VHB3 as similar as possible to the HB13 model so that parameter estimates of the two models would be comparable. Because this application results in three parameters (see below), we called this the VHB3 model. Here is the idea of VHB3. The participant first generates his or her FJ and then his or her HJ (i.e., Ernie’s judgment). The VHB3 model assumes that the following processes underlie HJs: The participant successfully retrieves his or her FJ and uses it as an HJ judgment with probability  $r'$ . By definition, this process will result in no hindsight bias ( $HJ = FJ$ ). In contrast, if the participant does not use the FJ to guide the HJ process (with probability  $1 - r'$ ), then two things can happen: (a) an FJ reconstruction that is unbiased by outcome knowledge (probability  $1 - b'$ ) or (b) an FJ reconstruction that is biased by outcome knowledge (probability  $b'$ ; see Figure 2).

Consider Case a, involving unbiased reconstruction, first. In this case, HJ will differ from FJ unsystematically, that is  $p(HJ > FJ) =$

$p(HJ < FJ) = 0.5$ . In Case b, involving biased reconstruction, two things can happen: confusing the current knowledge with the foresight knowledge (or the participant’s knowledge with Ernie’s knowledge) with probability  $c'$ —in this case, a maximum HJ judgment will occur—or HJ systematically deviates from FJ but does not match the current knowledge (probability  $1 - c'$ ). Thus, we have three parameters that have meanings very similar to the  $r_E$ ,  $b$ , and  $c$  parameters of the HB13 model:

$r'$ : the probability of successfully retrieving and using the FJ as the HJ.

$b'$ : the probability of being biased by outcome knowledge in generating the hindsight bias judgment, given that a participant did not use his or her FJ as HJ.

$c'$ : the probability of confusing current knowledge with foresight knowledge.

## Method

### Participants

There were 194 participants in total. Twelve groups of people participated: eighteen 3-year-olds ( $M = 3.68$  years,  $SD = .31$ ; 10 female), sixteen 4-year-olds ( $M = 4.53$  years,  $SD = .23$ ; 12 female), eighteen 5-year-olds ( $M = 5.46$  years,  $SD = .18$ ; 12 female), seventeen 6-year-olds ( $M = 6.36$  years,  $SD = .27$ ; nine female), thirteen 7-year-olds ( $M = 7.42$  years,  $SD = .30$ ; four female, eight male, one missing), fourteen 8-year-olds ( $M = 8.49$  years,  $SD = .32$ ; five female), nineteen 9-year-olds ( $M = 9.38$  years,  $SD = .26$ ; 10 female), thirteen 10-year-olds ( $M = 10.23$  years,  $SD = .26$ ; seven female), twelve 11-year-olds ( $M = 11.29$  years,  $SD = .33$ ; eight female), sixteen 13- to 15-year-olds ( $M = 13.47$  years,  $SD =$

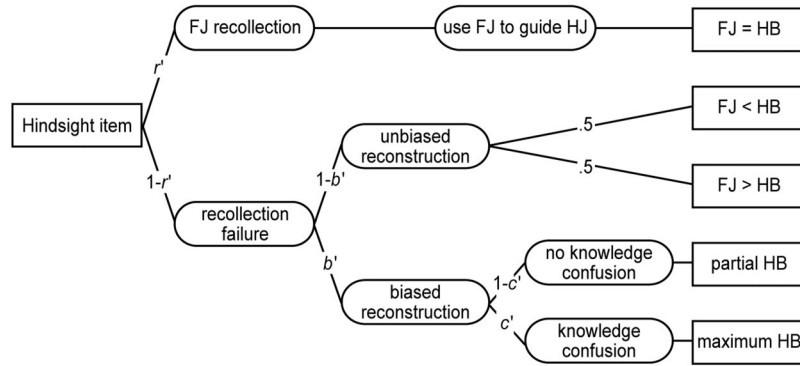


Figure 2. An illustration of the VHB3 model for analyzing visual HB data. Parameter  $r'$  denotes the probability of recalling one's FJ and using it as one's HJ. Parameter  $b'$  denotes the probability of a biased reconstruction given recollection failure. Parameter  $c'$  denotes the probability of confusing outcome knowledge with foresight knowledge given recollection failure. FJ = foresight judgment; HJ = hindsight judgment; HB = hindsight bias.

.51; 10 female), 16 younger adults ( $M = 19.50$  years,  $SD = 2.90$ ; age range: 18–29 years; 10 female), and 22 community-dwelling older adults ( $M = 84.59$  years,  $SD = 7.82$ ; age range: 61–95 years; 17 female). As in many adult human psychological studies, the majority of our participants were female; our findings, therefore, may be especially representative of females. Children and young adults came from schools in Seattle, Washington, while older adults came from independent-living retirement homes and community drop-in centers in and around Vancouver, British Columbia, Canada. Participants in all age groups were predominantly middle- and upper-middle class and English speaking. Four 3-year-olds, one 4-year-old, and one 5-year-old dropped out of the study early in the testing for a variety of reasons, including boredom and difficulty concentrating. These participants were replaced.

## Materials

**Verbal hindsight.** Using a task developed previously (Hell et al., 1988), we administered 20 questions (see the Appendix) designed to be educationally relevant and engaging to all age groups tested. Participants were told that the answers to all questions fell between one and 100. Extensive pilot testing on preschoolers revealed that older children could count up to 100 and understood the concept of numbers; however, some 3- and 4-year-olds had difficulty counting past 30, and for this reason, we are cautious about interpreting the results of this task for the youngest age groups tested.

In an OJ condition, participants answered all 20 questions. Later, during a surprise memory test, participants tried to recall their OJ answers (ROJ). For half the memory test items, participants learned the correct answers to the questions before trying to recall their original answers (experimental items). For the remaining items, participants simply tried to recall their original answers (control items). We used a single fixed order of questions; however, participants received one of four versions. In two versions, participants answered the odd-numbered questions before answering the even-numbered questions. In the remaining two versions, participants answered the even-numbered questions first. Experimental and control items were counterbalanced across these four

versions and within each age group. Order of presentation was identical on the FJ and HJ trials.

We calculated overall verbal hindsight bias using Pohl's (2007)  $\Delta z$  index, which is less susceptible to outliers than the Hell et al. (1988) index introduced previously. The idea here is that if a participant has hindsight bias, his or her recalled responses to experimental hindsight items will be closer to the correct answers than will his or her responses to control items. So, first one measures the shift in judgment between the hindsight and foresight conditions. This shift is computed using the two discrepancies  $|OJ - CJ|$  and  $|RJ - CJ|$ . If there is hindsight bias, then  $|RJ - CJ|$  will be the smaller of these in the case of experimental items. To compare across items that may have different natural scales, one simply divides all the discrepancies for each item by the standard deviation of the responses recorded for that item. We will indicate these scaled quantities using an overbar, that is,  $\overline{OJ}$ ,  $\overline{RJ}$ , and  $\overline{CJ}$ .<sup>1</sup> This scaling makes it possible to compare and aggregate the measurement of hindsight bias across all items in a sensible way. So, finally, we have, for the hindsight bias rating for each subject,

$$HB = (|\overline{OJ} - \overline{CJ}| - |\overline{RJ} - \overline{CJ}|)_e - (|\overline{OJ} - \overline{CJ}| - |\overline{RJ} - \overline{CJ}|)_c,$$

where the subscript  $e$  indicates the aggregation of a subject's experimental items and the subscript  $c$  indicates the aggregation of a subject's control items. The aggregation is performed by taking the median; no other outlier rejection is employed. This yields a positive number if the average shift, at recall, toward the correct answer is more pronounced for experimental items than for control items, which is the essence of hindsight bias.

**Computer hindsight.** We adapted a task developed by Bernstein et al. (2004) and Harley, Carlsen, and Loftus (2004). Stimuli consisted of 20 line drawings of common objects (e.g., apple, bed, flower). We scaled pictures of each object to fit within a  $245 \times$

<sup>1</sup> Pohl (2007) proposed  $z$ -transformed values as scaled quantities. However, the hindsight bias formula presented here is equivalent to Pohl's  $\Delta z$  measure because the means cancel out of the  $z$ -score differences.

245-pixel square on a Macintosh G4 Power Book or laptop computer. We then degraded each object by blurring it. Blurring was accomplished for each object by spatial low-pass filtering. Degree of blur was determined by the cutoff frequency of the low-pass filter and, for technical reasons, is characterized as an effective distance (measured in feet; for details, see Bernstein, Loftus, & Meltzoff, 2005; G. R. Loftus, 2002). For each object, we created 30 increasingly degraded images so that the differences between successive degraded images were roughly equal perceptually. During the baseline phase of the experiment, each object clarified from fully degraded to fully clear (see Figure 3).

A trial consisted of a single clarifying object along with the participant's associated response. The participant's task on half the trials (FJ) was to identify the object as soon as possible. At the outset of each FJ trial, participants were unaware of the object's identity. For each object, the 30 blurred images appeared, in order, from most to least blurred, at a rate of 500 ms per image. To the observer, it looked as if the object were slowly becoming clearer over time. Participants' task on the remaining trials (HJ) was to view the same objects that they had identified in the FJ condition and estimate when a same-age peer could identify the objects. Preschoolers (ages 3, 4, and 5 years) estimated for the Sesame Street puppet, Ernie, while all other participants estimated for an imaginary character named Ernie.<sup>2</sup> For the HJ condition, all participants learned that

Ernie is [x] years old, just like you, and he's just as smart as you. Ernie's never seen any of these pictures or played this game. This time, I want you to watch and tell me when Ernie can guess what's in the picture.

At the outset of each HJ trial with Ernie, the object appeared in full clarity while Ernie remained out of view. Children were reminded that Ernie could not see the object. This kind of puppet procedure has been used by many developmental researchers (see, e.g., Miller, 2007). The fully clarified object then disappeared from the screen, and the experimenter placed Ernie directly in front of participants and the computer screen. The experimenter pressed a key, and an object appeared maximally blurred, along with its name. The experimenter explained that the participant should call for the clarification to be stopped at the point when Ernie would be

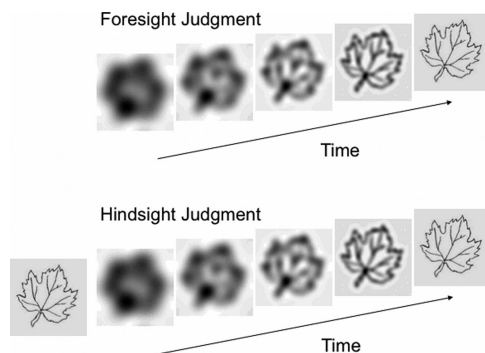


Figure 3. Example of computer hindsight task. In the foresight judgment condition, participants try to identify the object as it clarifies on a computer screen. In the hindsight judgment condition, participants first see the clear object and then indicate when someone else will be able to identify the object as it clarifies on a computer.

able to identify the object. The picture began clarifying, again at 500 ms per clarity level, and the experimenter stopped the clarification (with the space bar) when the participant reported that Ernie would now be able to identify the object. Immediately after the experimenter hit the space bar, the object disappeared, the experimenter confirmed that the subject was ready to proceed, and then the experimenter pushed *g* to proceed to the next object. Participants completed all 20 FJ trials before completing the 20 HJ trials.<sup>3</sup>

Participants sat facing the computer, with an experimenter seated directly beside them. The experimenter told participants that if they or Ernie did not know the object's identity or that it did not look like anything, they could say "don't know" or "nothing." The experimenter typed participants' responses in the foresight condition. Four different object orders were used, counterbalanced within each age group. Order of presentation was identical on the FJ and HJ trials.

We calculated visual hindsight bias for each participant as the difference between the mean identification point in the HJ condition and the mean identification point in the FJ condition. For example, a participant who identified objects on average at an effective distance of 300 feet in the HJ condition and 200 feet in the FJ condition would obtain a hindsight score of  $300 - 200 = 100$ . Thus, the greater the resulting value, the more degraded the image was, on average, in the HJ condition than in the FJ condition. This signifies visual hindsight bias. Note that the way in which we calculate visual hindsight bias differs from previous work in that here, we calculate a difference score, whereas, in prior work, researchers have calculated visual hindsight bias as a ratio score (e.g., Harley et al., 2004). After carefully examining the distribution of responses in the two visual hindsight tasks in the current experiment, we determined that a difference score better captured the nature of visual hindsight bias than did a ratio score. HJs were well correlated with FJs (indicating competence at the task), but the two distributions (hindsight and foresight) differed from one another by a constant value, which we interpret as hindsight bias.

**Real object hindsight.** We used the task developed by Bernstein et al. (2007) to assess visual hindsight bias using 3-D objects. We used four real objects, measuring up to 5 in. (12.7 cm) long and 5 in. high: red car, yellow glasses, green and white airplane, orange and brown alligator. The experimenter placed each object on a platform that was at eye level to the participant. This platform

<sup>2</sup> In pilot studies, we had also attempted to run several different types of memory design in which, during the HJ trials, participants tried to recall the precise point at which they identified each of the objects during the FJ. Unfortunately, although the memory design works well with adults (Bernstein & Harley, 2007; Harley et al., 2004), we could not get young preschoolers to understand this task. We, therefore, used a quasi-hypothetical hindsight design instead of a memory design.

<sup>3</sup> We have used different procedures in different studies, some published (see Bernstein et al., 2004, Experiment 1) and some unpublished. In some procedures, we used different items in the FJ and HJ conditions. In the past, we also have counterbalanced the order of presentation between FJs and HJs. These different procedures all yield robust hindsight bias. We opted to use the present within-participant, within-item procedure for its high statistical power but acknowledge that some of our conclusions may reflect this particular aspect of our experimental design.

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 were placed in a three-ring binder, which sat atop the box. The sheets hung in front of the toy, obscuring its appearance. Each sheet contained a different laminated transparency. Each transparency contained a unique set of black dots that covered 5% of the surface area of the sheet. Note that objects are easily identified behind 10 blank transparencies, but behind the 10 sheets used here, the object is completely occluded and impossible to identify.

The experimenter introduced this task as the “hide it game” (see Bernstein et al., 2007, for details). As in the computer hindsight task, participants first completed the FJ condition, attempting to identify each of the four objects as it gradually became clearer. 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 HJ condition, participants saw the object at the start of each trial. The experimenter then covered the object with all 10 filter sheets and asked participants to indicate when Ernie could identify the object. The experimenter removed one sheet at a time and asked participants, “What does it look like to Ernie now?” The experimenter recorded the point at which participants reported that Ernie could identify the object, by recording the number of sheets that had been turned. After the participant indicated this identification point for Ernie, the next trial began. Participants completed the four FJ trials before completing the four HJ trials. We used a single fixed order of presentation for the four objects. Order of presentation was identical on the FJ and HJ trials.

Hindsight bias was calculated as the difference between the mean identification point in the FJ condition and the mean identification point in the HJ condition. For example, a participant who identified objects on average at the eighth filter screen in the FJ condition and the fifth filter screen in the HJ condition would obtain a hindsight score of  $8 - 5 = 3$ . As in the computer hindsight task, the greater the resulting value, the more hindsight bias the participant exhibited.

## Procedure

One female or male experimenter tested participants individually in a single 1-hr session. Child and younger adult participants were tested in a quiet university laboratory or in a quiet room in their school, while older adults were tested in a quiet room either where they lived or in a community drop-in center. The testing session began with the OJ condition of the verbal hindsight task. This was followed by the FJ identification condition of the computer hindsight task, and then the FJ condition of the real object hindsight task. Next, participants completed the HJ condition of the computer hindsight task, followed by the HJ condition of the real object hindsight task. Finally, before starting the ROJ condition of the verbal hindsight task, older children (ages 10 years and up) and adults were asked to count backwards by threes from 300 until they reached zero. Preschoolers and younger children were asked to count to 100. The reason for this counting task was to create interference by making it harder for participants to recall their baseline answers in the verbal hindsight task. Participants were never told during the baseline conditions (OJ and FJ) of the

three hindsight tasks that they would later be tested on the questions or objects. Children received a gift for participating, and their parents received \$10. Younger adults received course credit, and older adults received \$10 for participating.

## Results

### Hindsight Bias Across the Life Span

As found previously (Bernstein et al., 2007), the two visual hindsight tasks correlated well ( $r = .44, p < .001$ ). The verbal hindsight task  $\Delta z$  measure correlated significantly with the real object hindsight task ( $r = .26, p < .001$ ), but not with the computer hindsight task ( $r = .01, p > .10$ ).

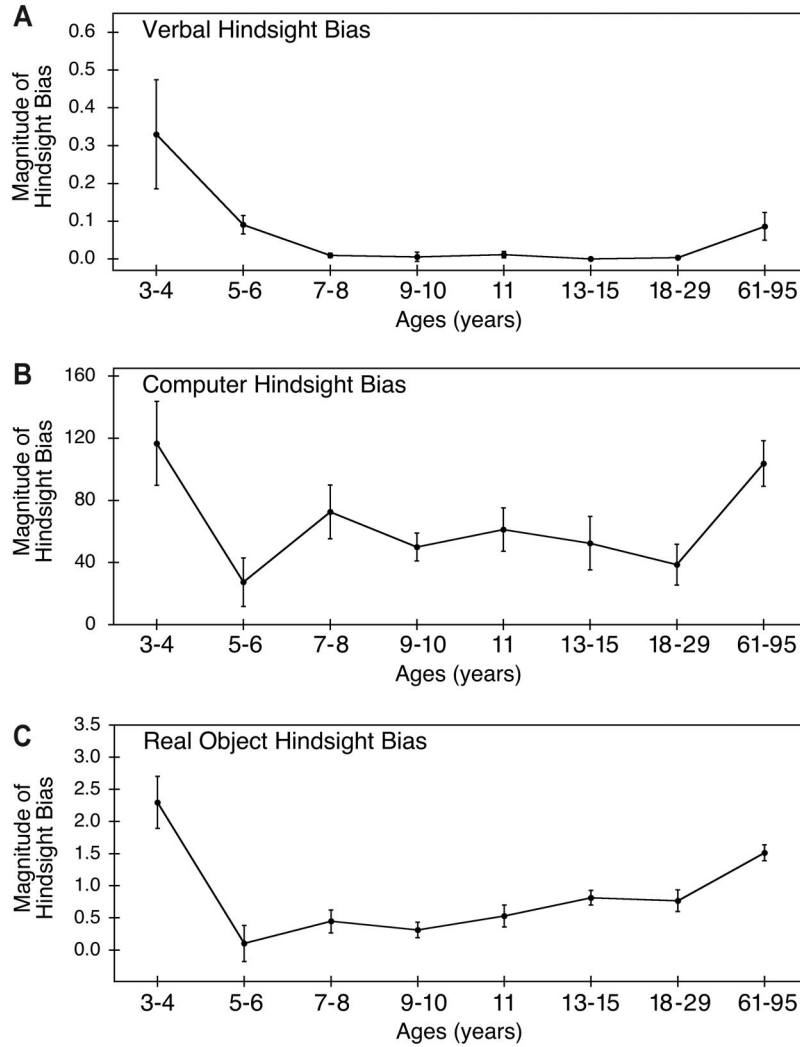
Figure 4 presents the results of our three hindsight tasks. For ease of presentation and to increase power of statistical tests, we combined ages as follows for the younger age groups: 3–4 years ( $M = 3.91$ ), 5–6 years ( $M = 5.89$ ), 7–8 years ( $M = 7.97$ ), and 9–10 years ( $M = 9.72$ ). Several points are worth noting about Figure 4. First, as can be seen in all three tasks, all age groups exhibited significant or near-significant overall verbal and visual hindsight bias. Second, in both visual hindsight tasks (see Figures 4B and 4C), hindsight bias declined during the preschool years from ages 3 to 6. Third, in all three tasks, hindsight bias followed a U-shaped function across age, with preschoolers and older adults exhibiting more hindsight bias than all other age groups and remarkable consistency in the magnitude of hindsight bias from age 6 to 30 years.

We conducted one-way analyses of variance (ANOVAs) on each of our three tasks, using our age-aggregated groups as the independent variable. The verbal hindsight task yielded a significant age difference in verbal hindsight bias,  $F(7, 186) = 3.21, p = .003$ . So too did the computer and real object visual hindsight tasks,  $F(7, 178) = 3.52, p = .001$ , and  $F(7, 192) = 9.58, p < .001$ , respectively. Next, we tested whether hindsight bias declined significantly with age in preschoolers. If hindsight bias relates to false-belief understanding, as has been found previously (Bernstein et al., 2007), we would expect to see a significant decline in hindsight bias from ages 3 to 5 years. The reason for this is that false-belief errors sharply decline between 3 and 5 years of age (e.g., Wellman, Cross, & Watson, 2001).

For this analysis, we kept our 3-, 4-, and 5-year-old age groups separate and conducted a one-way ANOVA on each of our three hindsight tasks. The verbal hindsight task yielded a significant age effect,  $F(2, 47) = 3.36, p = .044$ . The computer visual hindsight task yielded a marginally significant decline in the magnitude of hindsight bias,  $F(2, 46) = 3.39, p = .097$ . The real object visual hindsight task yielded a significant decline in the bias,  $F(2, 51) = 4.02, p = .024$ .

Post hoc comparisons yielded the following: For the verbal hindsight task, 3-year-olds ( $M = 0.07$ ) showed marginally less bias than 4-year-olds ( $M = 0.57$ ),  $t(29) = 1.81, p = .08$ , who showed more bias than 5-year-olds ( $M = 0.09$ ),  $t(31) = 2.19, p = .04$ . There was no difference between 3- and 5-year-olds. For the computer visual hindsight task, 3-year-olds ( $M = 116.12$ ) showed more bias than 5-year-olds ( $M = 31.41$ ),  $t(30) = 2.16, p = .04$ . While 4-year-olds ( $M = 117.17$ ) showed marginally more bias than 5-year-olds,  $t(31) = 1.90, p = .07$ , 3- and 4-year-olds did not differ significantly. For the real object visual hindsight task, again,





*Figure 4.* Magnitude of hindsight bias as a function of age groups. Panel A depicts Pohl's  $\Delta z$  index (described in detail in the text) for the verbal hindsight task. Panel B depicts the mean difference in effective distance, feet, at identification, between hindsight and foresight judgments in the computer hindsight task. Panel C depicts the mean difference in filter screen number between hindsight and foresight identifications in the real object hindsight task. Error bars show standard errors.

3-year-olds ( $M = 2.67$ ) showed more bias than 5-year-olds ( $M = 0.71$ ),  $t(34) = 2.93$ ,  $p = .006$ . While 4-year-olds ( $M = 1.88$ ) showed marginally more bias than 5-year-olds,  $t(32) = 1.80$ ,  $p = .08$ , 3- and 4-year-olds did not differ significantly. The developmental decline in visual hindsight bias from ages 3 to 5 years replicates the data pattern found in Experiment 2 in Bernstein et al. (2007). However, this decline does not generalize to the verbal hindsight task used in the current study, probably because the verbal task requires a deeper understanding of number magnitudes than was possible for our preschoolers.

### Multinomial Model-Based Analyses

**Verbal hindsight bias.** As is usual in HB13 analyses (Erdfelder & Buchner, 1998), we excluded cases in which OJ = CJ. This occurred on approximately 4.5% of trials. We then calculated

the frequencies of all 20 OJ-CJ-ROJ rank orders (10 for control and 10 for experimental items) across participants within each of our eight age groupings. With these raw data, we performed HB13 ML parameter estimation using the multiTree computer program of Moshagen (2010). We then evaluated the fit of the model at the significance level  $\alpha = .05$  using the likelihood ratio chi-square statistic  $G^2$ , which has five degrees of freedom for each of the eight age groups. As in other MPT analyses, data sets including categories with zero frequencies were incremented by 1 in all categories. Comparisons with other reasonable treatments of zero frequencies indicated that the effect of this increment on the ML parameter estimates is negligible.

Aggregating across all 20 items from our verbal hindsight task (see the Appendix), we found good or acceptable model fit for seven of our eight age groups,  $G^2(5) \leq 10.28$ ,  $p \geq .067$ . For the

5- to 6-year-olds, in contrast, the HB13 model failed to fit the data,  $G^2(5) = 19.07$ ,  $p = .002$ . Inspection of raw frequencies showed that three items (5, 7, and 13) were primarily responsible for the misfit. For these items, the relative frequencies of perfect recollections (i.e., ROJ = OJ) and source confusions (i.e., ROJ = CJ) differed greatly between participants who underestimated (OJ < CJ) versus overestimated (OJ > CJ) the CJ initially. Therefore, these items seem inappropriate for the HB13 model because the model presumes that recollection parameters ( $r_C$ ,  $r_E$ ) and source confusion parameters ( $c$ ) do not depend on whether the CJ is under- or overestimated initially. We thus followed Bayen et al. (2006) and eliminated the misfitting items from the data set. As expected, using the remaining 17 questions, we obtained good or acceptable HB13 model fits for each of our age groups, all  $G^2(5) \leq 10.59$ ,  $p \geq .060$ . The results reported here are based on these analyses.<sup>4</sup>

ML parameter estimates and standard errors of the recollection probabilities  $r_C$  and  $r_E$ , reconstruction bias  $b$ , and the source confusion parameter  $c$  are summarized in Table 1, separately for our different age groups. As expected, we found statistically significant age differences in recollection parameters  $r_C$ ,  $G^2(7) = 89.60$ ,  $p < .001$ , and  $r_E$ ,  $G^2(7) = 82.89$ ,  $p < .001$ . Younger children and older adults showed lower recollection rates compared with older children and younger adults for both the control and experimental items. This is consistent with an inverted U-shaped function depicting the development of memory ability across the life span (Zelazo, Craik, & Booth, 2004).

As can be seen in Table 1, even though descriptively we observed better recall of control items than experimental items for all age groups ( $r_C > r_E$ ), this difference was not significant for any of the age groups,  $G^2(1) \leq 3.67$ ,  $p \geq .055$ . Recollection bias  $r_C - r_E$ , that is, the deterioration in OJ recall memory induced by processing the CJ, is quite small (about .05) and appears to be almost constant across the life span. This is consistent with what Erdfelder, Brandt, and Bröder (2007) reported for recollection bias in young adults using a within-subject design such as the one used here.

As also can be seen in Table 1, all but three (9- to 10-year-olds, 11- to 13-year-olds, and 18- to 29-year-olds),  $G^2(1) \leq 3.40$ ,  $p \geq .065$ , of our age groups showed a significant reconstruction bias; that is, model parameter  $b$ , the probability of being biased by CJ knowledge when reconstructing the forgotten OJ, was significantly larger than zero for most age groups, all five  $G^2(1) \geq 4.53$ ,  $p \leq .033$ . Additionally, comparing across ages, reconstruction bias appears to be almost constant from preschool to younger adulthood (about  $b = .30$ ),  $G^2(6) = 3.69$ ,  $p = .72$ , and then increases in old age ( $b = .53$ ). This replicates Bayen et al. (2006, Experiment 1), although the increase in  $b$  between younger and older age groups was not significant in our data,  $G^2(1) = 1.78$ ,  $p = .18$ .

Finally, as can be seen in Table 1, younger children (3- to 4-year-olds, 5- to 6-year-olds, and 7- to 8-year-olds) and older adults were the only age groups to show significant source confusion bias; that is, model parameter  $c$ , the probability of confusing the CJ with the OJ given a biased reconstruction, was significantly above zero for these four age groups, all  $G^2(1) \geq 6.43$ ,  $p \leq .012$ , but not for the remaining groups of 9- to 29-year-olds, all  $G^2(1) \leq 2.94$ ,  $p \geq .086$ . More importantly, comparing across ages, the youngest age group tested here (3- to 4-year-olds) showed the most source confusion. Thus, although there is no clear evidence for a larger quantity  $b$  of being biased by outcome knowledge in preschoolers, there is strong evidence for reconstruction bias having a

different quality in preschoolers. Parameter  $c$ , which did not differ significantly between the age groups ranging from 5 to 95 years of age,  $G^2(6) = 6.04$ ,  $p = .42$ , was significantly higher in young preschoolers ( $c = .54$ ) compared to all other age groups,  $G^2(1) = 8.47$ ,  $p = .004$ .

**Visual hindsight bias.** We excluded cases from multinomial analyses as follows: With the FJ and HJ scales ranging from 1 (blurriest) to  $N$  (clearest), we omitted all cases with  $FJ < 3$  (because partial and maximum hindsight bias cannot be differentiated;  $FJ < 3$  occurred on 0.9% and 3.5% of computer and real object trials, respectively) and  $FJ = N$  (because reverse hindsight bias cannot occur;  $FJ = N$  occurred on 8.5% and 3.2% of computer and real object trials, respectively). As with HB13, we then calculated the frequencies of all four FJ–HJ pairings to which the model refers— $HJ = FJ$  (no hindsight bias),  $HJ > FJ$  (reverse hindsight bias),  $1 < HJ < FJ$  (partial hindsight bias), and  $HJ = 1$  (maximum hindsight bias)—across participants within each of our eight age groupings. With these raw data, we estimated the three model parameters with the multiTree computer program of Moshagen (2010) using the ML method. Because the number of model parameters (3) equals the number of free data categories ( $4 - 1 = 3$ ), the VHB3 model, as a saturated model, fits the observed rank-order frequencies perfectly ( $G^2 = 0$ ). Hence, goodness-of-fit statistics are not reported.

First, consider the computer visual hindsight task, illustrated in Table 2. Here, we see generally low and consistent scores across age groups in recollection parameter  $r'$ . Next, contrast this with recollection parameter  $r'$  in the real object hindsight task, also illustrated in Table 2. Here, we found higher recollection scores and statistically significant age differences in recollection parameter  $r'$ . As with our verbal hindsight results, preschoolers and older adults showed lower recollection rates compared with older children and younger adults. The most likely reason for the lower recollection scores in the computer visual hindsight task is that this task contained more than 10 times the number of images (30 levels of degradation  $\times$  20 objects shown = 600 images) than did the real object hindsight task (10 filter screens  $\times$  four objects = 40 images). If participants tried to use their own prior performance during the FJ phase to guide their estimates for Ernie during the HJ phase, they would be less likely to recall their FJ in the computer hindsight task than in the real object hindsight task.

As also can be seen in both the computer and real object hindsight tasks, illustrated in Table 2, nearly every age group showed a significant reconstruction bias; that is, model parameter  $b'$ , the probability of being biased by current knowledge given FJ recollection failure, was significantly above zero in 15 of 16 cases (5- to 6-year-olds showed no significant reconstruction bias on the real object hindsight task). Additionally, comparing across ages on the computer hindsight task (see Table 2), reconstruction bias appears to be almost constant from preschool to younger adulthood (about  $b' = .40$ ) and then increases in old age ( $b' = .69$ ). This replicates the data pattern that we observed for the verbal hindsight task and extends Bayen et al. (2006, Experiment 1) by showing that reconstruction bias is greater in older adults in visual hindsight bias, too.

<sup>4</sup>Raw frequency data based on these 17 items are available upon request.

Table 1

*HB13 Model Results for Verbal Hindsight Bias: Maximum-Likelihood Parameter Estimates (and Standard Errors) and  $G^2$  Goodness-of-Fit Statistics*

Parameter	Age in years							
	3–4	5–6	7–8	9–10	11	13–15	18–29	61–95
$r_C$	.31 (.03)	.44 (.03)	.55 (.04)	.56 (.03)	.69 (.05)	.73 (.04)	.70 (.04)	.47 (.04)
$r_E$	.30 (.03)	.36 (.03)	.54 (.03)	.50 (.03)	.55 (.05)	.70 (.04)	.64 (.04)	.42 (.04)
$r_C - r_E$	.01	.08	.01	.06	.14	.03	.06	.05
$b$	.32 (.11)	.42 (.09)	.39 (.12)	.13 (.14)	.38 (.19)	.39 (.17)	.31 (.22)	.53 (.12)
$c$	.54 (.18)	.22 (.07)	.29 (.07)	.07 (.17)	.02 (.11)	.30 (.16)	.03 (.13)	.12 (.05)
$G^2(5)$	5.00	10.59	6.74	4.15	4.08	4.00	4.68	2.26
$p$	.416	.060	.241	.529	.537	.550	.456	.812

*Note.*  $r_C$  = probability of original judgment recollection for control items (without feedback);  $r_E$  = probability of original judgment recollection for experimental items (with feedback);  $r_C - r_E$  = recollection bias induced by feedback presentation;  $b$  = reconstruction bias (the probability of a correct judgment-based reconstruction given recollection failure);  $c$  = probability of source confusion (reconstructing the correct judgment rather than the original judgment);  $p$  =  $p$  value.

When we examine reconstruction bias (parameter  $b'$ ) in the real object hindsight task (see Table 2), we see relatively and consistently high scores and no clear developmental pattern. Although older adults showed elevated reconstruction bias ( $b' = .87$ ), this value was not significantly higher than that of the younger adults or children. Because we observed similar result patterns (numerically larger reconstruction bias) among older adults on all three of our hindsight tasks, replicating Bayen et al. (2006), we conclude tentatively that older adults show significantly greater reconstruction bias than younger adults and children.

Finally, as can be seen in Table 2, young preschoolers (3- to 4-year-olds) were the only age group to show significant source confusion bias on both the computer and real object hindsight tasks; that is, model parameter  $c'$ , the probability of confusing their current knowledge of the object's identity with foresight knowledge, was significantly above zero. The result of such source confusion is that young preschoolers claimed that Ernie could identify the objects when there was not sufficient visual information present, and it was impossible for anyone to identify the objects without merely guessing (see Figure 3, most degraded image). Although 5- to 6-year-olds showed significant source confusion on the computer hindsight task ( $c' = .08$ , 95% CI [.01, .15]), they did not show significant source confusion on the real

objects task despite a larger ML estimate ( $c' = .18$ , 95% CI [-.02, .55]). More importantly, comparing across ages, the youngest age group tested here (3- to 4-year-olds) showed the most source confusion: Parameter  $c$  was significantly higher in preschoolers ( $c' = .33$  on the real object hindsight task) compared to all other age groups. Thus, as we observed in the verbal hindsight task, although there is no clear evidence for a larger quantity (parameter  $b$ ) of being biased by outcome knowledge in preschoolers, there is strong evidence for reconstruction bias having a different quality (parameter  $c$ ) in preschoolers.

## Discussion

This is the first study to trace the development of hindsight bias across the life span. Hindsight bias follows a U-shaped development, with preschoolers and older adults showing more bias than older children and younger adults (see Figure 4). Using MPT models to decompose visual and verbal hindsight bias, we found that young preschoolers exhibit a qualitative shift in their hindsight bias by confusing the outcome information with their original prediction and claiming that they or a naïve peer knew it all along. No other age group makes this kind of error consistently. Instead, older children and adults show hindsight bias by reconstructing

Table 2

*VHB3 Model Results for Visual Hindsight Bias: Maximum-Likelihood Parameter Estimates (and Standard Errors)*

Parameter	Age in years							
	3–4	5–6	7–8	9–10	11	13–15	18–29	61–95
Computer hindsight task								
$r'$	.06 (.01)	.07 (.01)	.06 (.01)	.08 (.01)	.07 (.02)	.07 (.02)	.07 (.02)	.05 (.01)
$b'$	.36 (.04)	.14 (.04)	.44 (.04)	.43 (.04)	.43 (.07)	.46 (.06)	.44 (.06)	.69 (.04)
$c'$	.04 (.01)	.08 (.04)	.01 (.01)	.01 (.01)	.01 (.01)	.01 (.01)	.01 (.01)	.00 (.00)
Real object hindsight task								
$r'$	.11 (.03)	.28 (.04)	.39 (.04)	.37 (.04)	.38 (.07)	.33 (.06)	.31 (.06)	.13 (.04)
$b'$	.67 (.07)	.12 (.10)	.54 (.10)	.47 (.10)	.64 (.14)	.90 (.06)	.78 (.09)	.87 (.06)
$c'$	.33 (.06)	.18 (.19)	.03 (.03)	.03 (.03)	.06 (.05)	.03 (.03)	.03 (.03)	.02 (.02)

*Note.*  $r'$  = probability of recollecting the foresight judgment and using it as the hindsight judgment;  $b'$  = reconstruction bias (probability of being biased by current knowledge given foresight judgment recollection failure);  $c'$  = probability of confusing current knowledge with foresight knowledge.

their original prediction and using the outcome information as a guide. Older adults tend to make this reconstructive error more than all other age groups. Thus, older adults show a different quantity of hindsight bias by more often forgetting their original prediction and reconstructing it using the outcome information.

We used both a memory design in which participants tried to recall their original answers to questions after learning the correct answers to those questions and a quasi-hypothetical design in which participants first identified degraded objects clarifying on a computer screen or behind a series of filter screens and then, armed with knowledge of the objects' identity, tried to estimate when a naïve peer would identify those objects. In both our verbal and visual tasks, all age groups exhibited hindsight bias by claiming that they themselves or a naïve peer knew more about the answers to the questions or about the identity of objects than the participants themselves actually knew. Thus, hindsight bias persists throughout life and follows a U-shaped function. The bias develops by age 3, tends to decline by age 5, then stabilizes but remains present in older children and younger adults before increasing in old age.

### MPT Models of Hindsight Bias in the Current Experiment

To explore the mechanisms underlying the U-shaped developmental pattern in hindsight bias that we observed, we conducted MPT models of our verbal and visual hindsight bias data. For our verbal hindsight task, we used the HB13 model previously developed by Erdfelder and Buchner (1998). For our two visual hindsight tasks, we developed a new VHB3 model fashioned after the HB13 model. These MPT models permitted us to estimate the relative strength of various theoretical constructs in our different age groups. In particular, we were interested in four constructs that are believed to contribute to hindsight bias: recollection ability, recollection bias, reconstruction bias, and source confusion.

With respect to recollection ability (parameters  $r_C$  and  $r_E$  in our verbal hindsight task and  $r'$  in the real object hindsight bias task), we replicated a data pattern observed by Bayen et al. (2006): Older adults showed lower recollection rates compared with younger adults for both the control and experimental items. In the present experiment, we also extended this finding to encompass the life span. Replicating prior work (e.g., Zelazo et al., 2004), recollection ability followed an inverted U-shaped curve across the life span such that preschoolers and older adults showed poorer recollection than did older children and younger adults.

With respect to recollection bias (parameter  $r_C - r_E$ ), or the likelihood of failing to recall one's original prediction after learning the outcome to an event, we again replicated Bayen et al. (2006): Recollection bias did not differ between older and younger adults. We also extended this finding to encompass the life span by showing that recollection bias remained small and constant from preschool to older adulthood. Apparently, the mechanism of item-specific retroactive inhibition, assumed to underlie recollection biases in within-subjects hindsight bias designs (Erdfelder et al., 2007), shows no clear developmental trend across the life span. Hence, this mechanism cannot be equated simply with susceptibility to retroactive interference or lack of inhibition, both of which have been argued to follow U-shaped curves across the life span (Bayen et al., 2007; Hasher & Zacks, 1988).

With respect to reconstruction bias (parameters  $b$  and  $b'$ ), or the likelihood of using outcome knowledge to help reconstruct one's original prediction, our parameter  $b$  values replicated the parameter estimates reported by Bayen et al. (2006) in their Experiments 1 and 2: Older adults showed elevated reconstruction bias compared to younger adults. In the present work, we also extended this finding to encompass the life span by showing that older adults had greater reconstruction bias compared to all other age groups that we tested in all three of our hindsight tasks. However, this elevation was significant only in the computer hindsight task. Given that we observed this elevation in older adults' parameters  $b$  and  $b'$  in all three tasks and that Bayen et al. observed this difference too, we conclude that older adults show a stronger quantity of hindsight bias than all other age groups: Older adults are more often biased by outcome knowledge, as measured by  $b$  and  $b'$ .

Finally, with respect to source confusion (parameters  $c$  and  $c'$ ), or the likelihood of confusing the outcome with one's original prediction, our parameter  $c$  values replicated the parameter estimates reported in Experiment 2 in Bayen et al. (2006); however, here, the difference between older and younger adults was not significant. We note that Bayen et al. obtained a significant age difference in parameter  $c$  in only one of their three experiments. Given that we failed to observe this difference in any of our three hindsight tasks and that Bayen et al. observed it in only one out of three of their experiments, and because both we and Bayen et al. observed differences in reconstruction bias (parameter  $b$ ) between older and younger adults, we conclude that younger and older adults differ significantly in terms of the quantity rather than the quality of their reconstruction bias. Moreover, parameter  $c$  from our verbal hindsight task remained stable for all age groups that we tested except for young preschoolers (3- to 4-year-olds), who showed elevated source confusion. Young preschoolers were the only age group to demonstrate elevated source confusion on all three of our hindsight tasks. We therefore conclude that young preschoolers exhibit a stronger quality of reconstruction bias by showing more source confusion than older children and adults: Young preschoolers tend to confuse foresight knowledge with outcome knowledge, as measured by  $c$  and  $c'$ .

Having described the data pattern that we observed from preschool to old age, we next discuss what may underlie this data pattern. One possibility is that it relates to broader developments in children's thinking that are commonly referred to as ToM. On this account, when a preschooler without a mature ToM learns that an alligator has 76 teeth, she is likely to think that she always knew that fact because she finds it impossible to understand that she once held a false belief (Astington & Gopnik, 1991; Taylor, Esbensen, & Bennett, 1994). When an adult with a mature ToM learns this new fact, she is likely to think that she always knew that alligators had something close to, but not exactly, 76 teeth. The adult can grasp that she may have once held a false belief (or was in ignorance), but the new factual information now colors her judgment of what she once knew.

Our data pattern also points to strategic differences between older adults and preschoolers. Older adults seem to be using an anchoring and adjustment strategy whereby they anchor on the correct answer and adjust from it (see Hawkins & Hastie, 1990). Preschoolers, in contrast, seem to be using an accessibility or fluency strategy whereby they focus on the most accessible and fluent answer—the correct answer—and replace their original an-

swer with it (see also Atance, Bernstein, & Meltzoff, 2010; Bernstein et al., 2007; Sanna, Schwarz, & Small, 2002).

Another possible idea that relates to the observed data pattern involves developmental changes in executive function, in particular, inhibitory control and working memory. According to this notion, preschoolers and older adults do poorly on tasks that require inhibitory control and working memory because they have immature (preschoolers) or deteriorating (older adults) frontal lobes. Applied to the current data, preschoolers have trouble inhibiting the correct answer to a problem or identity of a visual stimulus. They, in turn, assume that they or a naïve peer knew it all along. Older adults also have trouble inhibiting this privileged knowledge, but they also realize that they did not know that an alligator has exactly 76 teeth. Thus, older adults can inhibit their true belief default better than preschoolers can (see Leslie, German, & Polizzi, 2005). Although executive function deficits could explain portions of the current data pattern, it is unclear exactly how immature frontal lobes would produce a qualitative shift in hindsight bias in preschoolers while deteriorating frontal lobes would produce a quantitative shift in hindsight bias in older adults.

There are several other factors that may have influenced our results. First, our testing session lasted 1 hr. Thus, fatigue in the youngest and oldest age groups may have influenced our findings. Second, our older adults were quite old, with a mean age of 84.59 years. Thus, the age differences in hindsight bias that we observed may partly reflect our definition of older adults. Despite this latter possibility, our MPT results closely match those of Bayen et al. (2006), whose older adult sample had a mean age of 72 years.

Finally, it is possible that our data pattern represents a continuous shift in hindsight bias over all ages that we tested. Although we observed a qualitative shift in preschoolers and young school-age children in terms of how they exhibited hindsight bias (through source memory confusion), it is possible that this qualitative shift in performance represents a quantitative difference. Cognitive development can appear discontinuous and qualitative when the underlying mechanisms are continuous and quantitative (e.g., Friend, 2004). Future work should explore whether the changes that we observed in the magnitude of hindsight bias reflect qualitative and/or quantitative differences. Future work should also explore the mechanisms underlying hindsight bias that we identified here and assess whether similar mechanisms apply to developmental changes in perspective taking and ToM more broadly. MPT models provide one elegant approach to exploring such mechanisms.

Why do humans exhibit hindsight bias in the first place? We, like others, believe that hindsight bias is a by-product of an adaptive process (Hawkins & Hastie, 1990; Hoffrage, Hertwig, & Gigerenzer, 2000): In most cases, it is better to update knowledge with the correct answer to a question or problem than it is to ignore outcome information. The downside of such automatic knowledge updating is that people tend to forget their original, naïve thoughts, views, and predictions. This can be problematic, especially in medico-legal contexts in which juries are asked to assign blame in medical malpractice cases (Harley, 2007). For example, knowing the outcome to a case (the patient died) has serious implications for assigning blame to the doctor responsible for care. Relatively quick and automatic knowledge updating makes for adaptive learning, but it can lead people astray. Forgetting one's original

naïve prediction is an inevitable consequence of having an automatic updating system.

Why would hindsight bias be in place by age 3 and persist throughout life? The reason is that hindsight bias is an unintended by-product of an otherwise adaptive process for selecting and processing information as people try to make sense of the events that they experience. This process is so efficient that it often prevents people from reconstructing the causal and temporal chain of events leading to a particular outcome, rendering the outcome inevitable or more likely (Fessel, Epstude, & Roese, 2009; Henriksen & Kaplan, 2003). The fact that preschoolers and older adults exhibit more hindsight bias than school-age children and younger adults indicates that people may improve their ability to ignore privileged information for the majority of their lives but that they never perfect this ability.

## References

- 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. doi:10.1037/0021-9010.66.2.252
- Astington, J. W., & Gopnik, A. (1991). Theoretical explanations of children's understanding of the mind. *British Journal of Developmental Psychology, 9*, 7–31.
- Atance, C. M., Bernstein, D. M., & Meltzoff, A. N. (2010). Thinking about false belief: It's not just what children say, but how long it takes them to say it. *Cognition, 116*, 297–301. doi:10.1016/j.cognition.2010.05.008
- Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology, 23*, 611–626. doi:10.1037/0012-1649.23.5.611
- Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6*, 57–86.
- Bayen, U. J., Erdfelder, E., Bearden, J. N., & Lozito, J. P. (2006). The interplay of memory and judgment processes in effects of aging on hindsight bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 1003–1018. doi:10.1037/0278-7393.32.5.1003
- Bayen, U. J., Pohl, R. F., Erdfelder, E., & Auer, T. S. (2007). Hindsight bias across the life span. *Social Cognition, 25*, 83–97. doi:10.1521/soco.2007.25.1.83
- 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. doi:10.1111/j.0963-7214.2004.00663.x
- Bernstein, D. M., Atance, C., Meltzoff, A. N., & Loftus, G. R. (2007). Hindsight bias and developing theories of mind. *Child Development, 78*, 1374–1394. doi:10.1111/j.1467-8624.2007.01071.x
- Bernstein, D. M., & Harley, E. M. (2007). Fluency misattribution and visual hindsight bias. *Memory, 15*, 548–560. doi:10.1080/09658210701390701
- Bernstein, D. M., Loftus, G. R., & Meltzoff, A. N. (2005). Object identification in preschool children and adults. *Developmental Science, 8*, 151–161. doi:10.1111/j.1467-7687.2005.00402.x
- Birch, S. A. J., & Bernstein, D. M. (2007). What can children tell us about hindsight bias: A fundamental constraint on perspective-taking? *Social Cognition, 25*, 98–113. doi:10.1521/soco.2007.25.1.98
- Birch, S. A. J., & Bloom, P. (2003). Children are cursed: An asymmetric bias in mental-state attribution. *Psychological Science, 14*, 283–286. doi:10.1111/1467-9280.03436
- Blank, H., Fischer, V., & Erdfelder, E. (2003). Hindsight bias in political elections. *Memory, 11*, 491–504. doi:10.1080/09658210244000513
- Blank, H., & Nestler, S. (2007). Cognitive process models of hindsight bias. *Social Cognition, 25*, 132–146. doi:10.1521/soco.2007.25.1.132

- Blank, H., Nestler, S., von Collani, G., & Fischer, V. (2008). How many hindsight biases are there? *Cognition*, *106*, 1408–1440. doi:10.1016/j.cognition.2007.07.007
- Dehn, D. M., & Erdfelder, E. (1998). What kind of bias is hindsight bias? *Psychological Research*, *61*, 135–146. doi:10.1007/s004260050020
- Erdfelder, E., Auer, T. S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models: A review of the literature. *Zeitschrift für Psychologie/Journal of Psychology*, *217*, 108–124.
- Erdfelder, E., Brandt, M., & Bröder, A. (2007). Recollection biases in hindsight judgments. *Social Cognition*, *25*, 114–131. doi:10.1521/soco.2007.25.1.114
- Erdfelder, E., & Buchner, A. (1998). Decomposing the hindsight bias: A multinomial processing tree model for separating recollection and reconstruction in hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 387–414. doi:10.1037/0278-7393.24.2.387
- Fessel, F., Epstude, K., & Roese, N. J. (2009). Hindsight bias redefined: It's about time. *Organizational Behavior and Human Decision Processes*, *110*, 56–64. doi:10.1016/j.obhdp.2009.07.001
- Fischhoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 288–299. doi:10.1037/0096-1523.1.3.288
- Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 424–444). Cambridge, England: Cambridge University Press.
- Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology*, *50*, 21–45. doi:10.1146/annurev.psych.50.1.21
- Friend, M. (2004). On the implications of curvilinear trajectories for cognitive development. *Journal of Cognition and Development*, *5*, 103–108. doi:10.1207/s15327647jcd0501\_8
- Harley, E. M. (2007). Hindsight bias in legal decision making. *Social Cognition*, *25*, 48–63. doi:10.1521/soco.2007.25.1.48
- 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, and Cognition*, *30*, 960–968. doi:10.1037/0278-7393.30.5.960
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). New York, NY: Academic Press.
- Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, *107*, 311–327. doi:10.1037/0033-2909.107.3.311
- Hell, W., Gigerenzer, G., Gauggel, S., Mall, M., & Müller, M. (1988). Hindsight bias: An interaction of automatic and motivational factors? *Memory & Cognition*, *16*, 533–538.
- Henriksen, K., & Kaplan, H. (2003). Hindsight bias, outcome knowledge, and adaptive learning. *Quality and Safety in Health Care*, *12*, ii46–ii50. doi:10.1136/qhc.12.suppl\_2.ii46
- 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. doi:10.1037/0278-7393.26.3.566
- Hu, X., & Batchelder, W. H. (1994). The statistical analysis of general processing tree models with the EM algorithm. *Psychometrika*, *59*, 21–47. doi:10.1007/BF02294263
- Leary, M. R. (1981). The distorted nature of hindsight. *Journal of Social Psychology*, *115*, 25–29. doi:10.1080/00224545.1981.9711984
- Leary, M. R. (1982). Hindsight distortion and the 1980 Presidential election. *Personality and Social Psychology Bulletin*, *8*, 257–263. doi:10.1177/0146167282082012
- Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology*, *50*, 45–85. doi:10.1016/j.cogpsych.2004.06.002
- Loftus, E. F., & Loftus, G. R. (1980). On the permanence of stored information in the human brain. *American Psychologist*, *35*, 409–420. doi:10.1037/0003-066X.35.5.409
- Loftus, G. R. (2002). Analysis, interpretation, and visual presentation of data. In H. Pashler & J. Wixted (Eds.), *Stevens' handbook of experimental psychology: Vol. 4. Methodology in experimental psychology* (3rd ed., pp. 339–390). New York, NY: Wiley.
- Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. (2009, July 17). Foundations for a new science of learning. *Science*, *325*, 284–288. doi:10.1126/science.1175626
- Miller, S. A. (2007). *Developmental research methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Moshagen, M. (2010). multiTree: A computer program for the analysis of multinomial processing tree models. *Behavior Research Methods*, *42*, 42–54. doi:10.3758/BRM.42.1.42
- MSNBC News. (2006, February 10). Brown blames homeland for Katrina response. Retrieved from <http://www.msnbc.msn.com/id/11255379/>
- National Commission on Terrorist Acts Upon the United States. (2004). *The 9/11 Commission report*. Retrieved from <http://www.gpoaccess.gov/911/pdf/fullreport.pdf>
- Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.
- Pohl, R. F. (2007). Ways to assess hindsight bias. *Social Cognition*, *25*, 14–31. doi:10.1521/soco.2007.25.1.14
- Pohl, R. F., Bayen, U. J., & Martin, C. (2010). A multiprocess account of hindsight bias in children. *Developmental Psychology*, *46*, 1268–1282. doi:10.1037/a0020209
- Pohl, R. F., Bender, M., & Lachmann, G. (2002). Hindsight bias around the world. *Experimental Psychology*, *49*, 270–282. doi:10.1026/1618-3169.49.4.270
- 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.
- 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. doi:10.2307/1131282
- Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.
- 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. doi:10.1111/1467-8624.00304
- Wood, G. (1978). The knew-it-all-along effect. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 345–353. doi:10.1037/0096-1523.4.2.345
- Zelazo, P. D., Craik, F. I. M., & Booth, L. (2004). Executive function across the life span. *Acta Psychologica*, *115*, 167–183. doi:10.1016/j.actpsy.2003.12.005
- 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. doi:10.1006/obhd.1995.1093

(Appendix follows)

**Appendix****List of Questions (and Correct Answers) Used in the Verbal Hindsight Task**

1. How many inches across is the eye of a giant squid? (15)
  2. How many neck bones does a giraffe have? (7)
  3. How many seats are there on a school bus? (24)
  4. How many minutes does it take light from the sun to reach Earth? (8)
  5. How many legs does a lobster have? (10)
  6. How many provinces does Canada have? (10)
  7. How many days can a cockroach live without a head? (9)
  8. How many years can a parakeet live? (15)
  9. How many teeth does an alligator have? (76)
  10. How many miles per hour can a hippo run? (20)
  11. How many countries are in South America? (13)
  12. How many muscles does it take to frown? (43)
  13. How many teeth does a mosquito have? (47)
  14. How many pounds is a sperm whale's brain? (20)
  15. How many hours does a lion sleep in a day? (20)
  16. How many countries are in Europe? (45)
  17. How many moons does the planet Saturn have? (46)
  18. How many weeks are female dogs pregnant? (9)
  19. How many countries are there in Africa? (53)
  20. How many feet can a kangaroo jump in one leap? (30)
- Received March 9, 2010  
Revision received September 18, 2010  
Accepted September 29, 2010 ■