The Effect of the Number of A Trials on Performance on the A-Not-B Task

By: Stuart Marcovitch, Philip David Zelazo, Mark A. Schmuckler

Marcovitch, S., Zelazo, P. D., & Schmuckler, M. A. (2002). The effect of number of A trials on performance on the A-not-B task. *Infancy, 3*, 519-529.

Abstract:
The A-not-B error (Piaget, 1954), which occurs when infants search perseveratively on reversal trials in a delayed-response task, is one of the most widely studied phenomena in developmental psychology. Nonetheless, the effect of A-trial experience on the probability and magnitude of this error remains unclear. In this study, 9-month-old infants were tested at location A until they searched correctly on 1, 6, or 11 A trials. Results revealed an effect of A trials on the proportion of infants who erred on the first B trial, and on the number of errors prior to a correct search at B (i.e., the error run). These effects were asymptotic, or U-shaped, consistent with a dual-process model according to which A-trial experience increases habit strength but also provides opportunities for reflection on task structure.

Article:
The A-not-B error, first described by Piaget (1954), is one of the most widely studied phenomena in developmental psychology (for reviews, see Gratch, 1975; Harris, 1986, 1987; Marcovitch & Zelazo, 1999; Schuberth, 1983; Wellman, Cross, & Bartsch, 1986). Although Piaget (1954) argued that the error was indicative of a transitional phase in the understanding of object permanence, subsequent researchers have linked the error to shortcomings in several abilities, including spatial representation (Bremner & Bryant, 1977; Butterworth, 1977), memory (Bjork & Cummings, 1984; Munakata, 1997), memory plus inhibition (Diamond, Cruttenden, & Neiderman, 1994), strategy selection (Wellman et al., 1986), appropriate organization of internal dynamics (Smith, Thelen, Titzer, & McLin, 1999), and conscious control (Marcovitch & Zelazo, 1999).

The typical A-not-B search task comprises two (or more) hiding locations. An infant watches as an object is hidden at one location (A). Following a brief delay (e.g., 3 sec), the infant is permitted to search for the object. After one or more A trials, the object is then hidden conspicuously at a different location (B). Despite witnessing the hiding event at B, most 8- to 12-month-olds continue to search perseveratively at A on the first B trial, thereby committing the A-not-B error.

Several researchers have noted that the A-not-B task is an interference paradigm, and they have suggested that multiple psychological processes compete to determine search. For example, Diamond et al. (1994) proposed that on B trials infants have a prepotent tendency to reach to A, which competes with short-term (or working) memory of the object's current location (B). A correct search at B requires infants to inhibit this prepotent tendency. The dorsolateral prefrontal cortex is required for working memory and inhibition, and the likelihood of effective inhibition and robust working memory increases as this region of the prefrontal cortex develops.

Similarly, Munakata (1998) suggested that the A-not-B error occurs when latent memory traces exert a stronger influence on behavior than active memory traces. Like Munakata, but unlike Diamond et al. (1994), Marcovitch and Zelazo (1999) argued that a separate inhibition mechanism is unnecessary to explain the A-not-B error. Instead, these authors showed mathematically that the combined influence of both a response-based system and a conscious representational system successfully predicts the probability of searching at a particular location on a multilocation version of the task (Diamond et al., 1994). From a rather different perspective, dynamic systems researchers (Smith et al., 1999; Thelen, Schöner, Scheier, & Smith, 2001) have also claimed that search is jointly determined by a host of psychological processes, including previous motor behavior and attentional focus.

Despite their differences, these competing systems accounts of infant search all predict that performance on the A-not-B task will vary as a function of the number of A trials. Most of these theories (Diamond et al., 1994; Munakata, 1998; Smith et al., 1999; Thelen et al., 2001) predict that the likelihood of persistent search on B trials will increase monotonically as the number of A trials increases, a prediction consistent with basic learning theory (e.g., Hull, Felsinger, Gladstone, & Yamaguchi, 1947; Thorndike, 1921).

Marcovitch and Zelazo's (1999) model also predicts that performance on the A-not-B task will vary as a function of the number of A trials. However, this model not only allows for an asymptotic effect, but also indicates that the effect may be nonmonotonic, reflecting the opposing effects of A trials on the response-based and conscious representational systems. More precisely, although increases in A trials will increase habit strength (mediated by the response-based system), they will also provide additional opportunities to understand the task structure (mediated by the conscious representational system). The combined influence of these two processes may produce a stable asymptotic effect at some point, or it may yield a nonmonotonic effect (e.g., a U-shaped effect). For example, the likelihood of perseveration may gradually increase with increases in the number of A trials, indicative of the increase in habit strength, but then decrease, indicative mainly of the increased opportunities to understand the task structure.[1]

At this point, it remains unclear whether the number of A trials has any influence at all on the A-not-B task. Several studies have failed to find an effect of the number of A trials across a narrow range (3 vs. 5 A trials, Butterworth, 1977; 1 vs. 2 A trials, Diamond, 1983; 2 vs. 5 A trials, Evans, 1973; 1 vs. 3 A trials, Sophian & Wellman, 1983). However, it is possible that small differences in the number of A trials might lead to small differences in performance that would be difficult to detect without a large sample size.

Wellman et al. (1986) conducted a meta-analysis to determine statistically which experimental variables predict perseverative behavior on the A-not-B task, and these authors failed to find an effect of the number of A trials. Although this null finding has been widely cited (e.g., Flavell, Miller, & Miller, 1993; Harris, 1989; Hofstadter & Reznick, 1996), there are at least four reasons to question its validity.

First, the dependent measure used in the meta-analysis was the proportion of infants who erred on the first B trial. A different measure, such as the *error run* (i.e., the number of errors prior to searching correctly at B), might be more sensitive in revealing an effect. Indeed, Landers (1971) reported a difference in the error run after 2 A trials versus 8 or 10 A trials. On the other hand, several studies failed to find an effect on the error run across a narrow range of number of A trials (Butterworth, 1977; Diamond, 1983; Evans, 1973; Sophian & Wellman, 1983).

Second, in the large majority of studies included in the meta-analysis by Wellman et al. (1986), three or fewer A trials were used. This limited range might have made it difficult to detect differences.

Third, Smith et al. (1999) recently reported that the probability of searching perseveratively on the first B trial increases as a function of the number of previous reaches to the A location, regardless of the number of rewarded A trials. To determine the number of reaches to A (i.e., to quantify levels of the independent variable), Smith et al. included reaches during training and spontaneous reaches that occurred during the testing session (see Diamond, 2001, for a critique of this method). Although this finding does not directly address the issue of the number of rewarded A trials in the standard A-not-B paradigm, it does encourage further research on this topic.

Finally, Marcovitch and Zelazo's (1999) recent meta-analysis of the A-not-B error revealed an effect of the number of A trials. Whereas Wellman et al. (1986) had 58 data points in their "standard multilocation analysis," Marcovitch and Zelazo (1999) were able to use 107. The earlier analysis may have failed to find an effect of the number of A trials in part because of the relatively small sample size. Note, however, that the updated meta-analysis suggests the existence of an effect of A trials, but it does not provide a direct test of this effect.

This study was designed to assess directly the effect of the number of A trials on perseverative responding in the A-not-B task using a wide range of number of A trials. Nine-month-old infants received A trials until they searched correctly 1, 6, or 11 times. All training was done at a neutral location so that the number of A trials only reflected the number of rewarded reaches to A during the task proper (cf. Smith et al., 1999). Unlike Smith et al. (1999), we required infants to search correctly the prescribed number of times before proceeding to the B trials. Moreover, as measures of perseveration, we used both the proportion of children who erred on the first B trial and the error run. Finally, because previous failures to find an effect of the number of A trials may be due to the use of relatively small sample sizes (e.g., Butterworth, 1977), we tested 28 infants per condition.

METHOD

Design and Overview

Infants were randomly assigned to a 1, 6, or 11 A-trials condition. Within each condition, for half the infants, A was on the left and B was on the right. The center location was used for the training trials only.

Participants

Eighty-four infants were tested within 2 weeks of their 9-month birthday (*M* age = 39.9 weeks, range = 37.7–42.1). Infants were recruited from a database that contained names of parents who previously expressed interest in participating in research. Although demographic information was not recorded, the cultural background of the infants appeared to reflect the cultural variety of the population of Toronto, Canada.

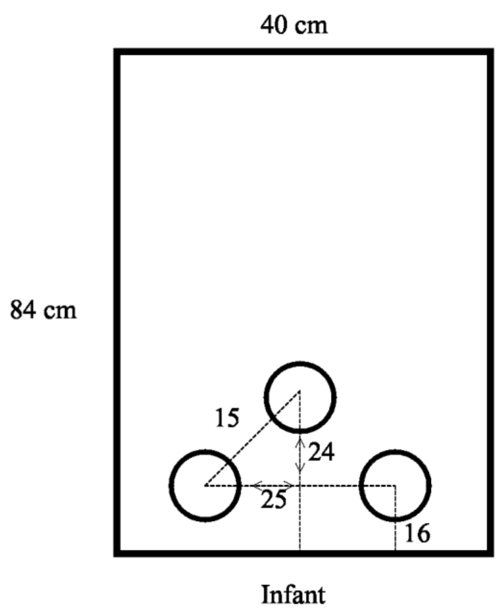

FIGURE 1 Sketch of task apparatus.

Sixty additional infants were eliminated from the final sample.[2] Of these, 39 refused to participate during the training phase, 19 failed to meet the A-trial criterion, and 2 were eliminated due to parental interference. Twelve of these infants had been assigned to receive 1 A trial, 24 to receive 6, and 24 to receive 11. Chi-squared analyses failed to reveal a relation between the probability of exclusion and assigned condition, $x^2(2, N= 60) = 4.80, p > .05$.

Materials

The apparatus consisted of a wooden box 84 cm long x 40 cm wide x 8 cm deep (see Figure 1). The box was tan with white sides, and on the top of the box on the side closest to the infant there were three wells (10 cm in diameter and 6 cm deep). The wells were arranged in a triangular array (see Figure 1 for exact distances). Each well was covered with a square patch of black cloth (14 cm x 14 cm). A horizontal slit, 8 cm wide, was made in each patch to allow access to the well, and Velcro was used to keep the slits closed. A piece of cardboard, 28.5 cm x 37.5 cm, was used to cover all three wells simultaneously (see Diamond et al., 1994).

The experimenter selected a toy from several small toys that he thought would be appealing to the infant. However, if the infant lost interest in the toy at any time during the procedure, a new one was selected.

Procedure

All testing was conducted by Stuart Marcovitch in one of two testing locations. The infant was seated on the parent's lap and positioned across a table from the experimenter. The parent was instructed to prevent the infant from searching during the hiding event and the delay, but not to interfere with the infant in any way.

Training. The training phase consisted of four trials presented in a fixed order (see Schmuckler & Tsang-Tong, 2000, for a similar procedure). On all training trials, the toy was hidden at the center location. If the infant did not retrieve the toy spontaneously within 30 sec, the infant was given the toy to play with (or the toy was revealed), and the trial was repeated until the toy was correctly retrieved or until the infant refused to participate further.

On the first training trial, the experimenter placed the toy in front of the center well, and the infant was encouraged to retrieve it immediately. On the second training trial, the experimenter placed the toy in the center well so that it was half visible, and again, the infant was encouraged to find it immediately. On the third training trial, the experimenter hid the toy completely in the center well, and the infant was encouraged to search for it immediately. On the fourth training trial, the toy was again hidden in the center well. However, before the infant was allowed to search, the experimenter covered the three wells simultaneously with the cardboard and counted 3 sec aloud. The cardboard was then removed and the infant was encouraged to search.

A trials. Following the training trials, A trials were administered in a manner identical to the last training trial, except that the toy was hidden at A. Infants were considered to have searched at a particular location if their hand entered the well (i.e., broke the horizontal plane). If the infant searched correctly, he or she was allowed to play with the toy for about 30 sec. Incorrect search or the failure to search within 30 sec resulted in readministration of the trial with no handling of the toy. A trials were administered until the infant searched the prescribed number of times correctly. If 20 A trials elapsed and the toy had still not been found the appropriate number of times, the session was stopped. However, if the infant searched correctly on the 20th A trial, testing continued until the criterion was reached or a subsequent error was committed.

B trials. B trials commenced immediately after the last A trial. These trials were identical to the A trials, except that the toy was hidden at B. Testing was concluded after the infant found the toy at B three times or refused to participate further.

RESULTS

In both the 6 and 11 A-trial conditions, the run of correct A trials could be consecutive, or disrupted by one or more incorrect searches (i.e., intrusions). The mean number of intrusions was 1.00 (SD = 2.02, range = 0–9) in the 6 A-trial condition and 1.25 (SD = 1.55, range = 0–5) in the 11 A-trial condition. This difference is not significant, $t(54) = 1.08$, $p > .1$.

The dependent measure most often cited in A-not-B research is the proportion of infants who search incorrectly on the first B trial (see Table 1). A chi-square analysis revealed a relation between the number of A trials and the proportion of infants who were incorrect on the first B trial, $x^2(2, N= 84) = 7.5 1$, $p < .05$. Follow-up tests comparing two groups directly indicated that infants who experienced 6 A trials were more likely to err on the first B trial than infants who experienced 1 A trial, $X^2(1, N = 56) = 7.49$, $p < .01$, and marginally more likely to err than infants who experienced 11 A trials, $x^2(1, N= 56) = 2.95$, $.05 < p < .1$. There was no difference between infants who experienced 1 A trial and infants who experienced 11 A trials, $X^2(2, N= 56) = 1.14$, $p > .1$.

TABLE 1
Proportion of Infants Who Searched Incorrectly on the First B Trial and Length of Error Run by Condition

	Number of A Trials		
	1	*6*	*11*
Proportion incorrect	.43 (12)	.79 (22)	.57 (16)
Error run[a]			
Median	0[b]	2[c]	1[d]
Range	0–6	0–18	0–14

[a]*ns* for this measure do not include infants who stopped before making one correct search on the B trials. [b]*n* = 27. [c]*n* = 28. [d]*n* = 26.

The error run (i.e., number of errors before the first correct search) was also analyzed. Table 1 reveals the median error run (and range) at each condition. The Kruskal–Wallis test showed an overall effect of condition on this measure, $x^2(2) = 6.56$, $p < .05$. Post hoc analyses revealed that infants who experienced 6 A trials had longer error runs than infants who experienced 1 A trial. However, there was no significant difference in the length of the error run between infants who experienced 11 A trials and infants in either of the other two groups.

Most infants (66 out of 84, 79%) eventually found the toy at B three times. However, 3 infants never found the toy, and 15 found the toy at least once but fewer than three times. These 18 infants were fairly evenly distributed across conditions (5 received 1 A trial, 6 received 6 A trials, and 7 received 11 A trials). Once infants searched correctly at the B location, they usually continued to search correctly on all subsequent trials. Considering only those infants who searched correctly at least once, 15 of 27 infants in the 1 A-trial condition, 19 of 28 in the 6 A-trial condition, and 14 of 26 in the 11 A-trial condition searched correctly on all subsequent trials. The difference among conditions was not significant, $x^2(2, N= 81) = 1.32$, $p > .1$.

DISCUSSION

Due largely to the results of the meta-analysis conducted by Wellman et al. (1986), many researchers concluded that the number of A trials has no effect on the A-not-B error, and this conclusion has served as an important constraint on theoretical explanations of infant search (e.g., Harris, 1989; Wellman et al., 1986). These results indicate that this conclusion was premature. Indeed, the results clearly show an effect of the number of A trials: There was a significant increase in errors between 1 and 6 A trials on both the proportion of infants who searched incorrectly on the first B trial and the error run.

Both Landers (197 1) and Butterworth (1977) failed to find an effect of A trials using the proportion of infants who searched incorrectly on the first B trial. However, the sample sizes in these studies were relatively small. Whereas Landers (1971) tested 14 infants per condition and Butterworth (1977) tested 8, we tested 28. It seems likely that previous studies, like the influential meta-analysis by Wellman et al. (1986), failed to find an effect of the number of A trials due to a lack of statistical power.

The increase in perseveration between the 1 and 6 A-trials conditions was expected on the basis of various competing systems theories of infant search (e.g., Diamond et al., 1994; Marcovitch & Zelazo, 1999; Munakata, 1998; Smith et al., 1999; Thelen et al., 2001), and it is consistent with classic learning theory. However, there was no evidence for a corresponding increase between the 6 and 11 A-trials conditions, and some evidence of a diminution. Thus, it is possible that the effect of A trials reaches an asymptote, which would be consistent with the Rescorla and Wagner (1972) model of classical conditioning. Alternatively, the effect may be U-shaped. Further work with additional levels of number of A trials will be required to distinguish between these options.

Asymptotic Pattern

A monotonic effect of the number of A trials is consistent with all competing systems theories of the A-not-B error. Although the claim has not been made explicitly, it is reasonable to assume that these theories would predict an asymptotic effect (and, indeed, it is a feature of Munakata's [1998] connectionist simulation).

If the pattern of behavior is asymptotic, then it can be accounted for by the influence of experience on a single process, such as habit strength (e.g., Rescorla & Wagner, 1972). One possible asymptotic function was suggested by Marcovitch and Zelazo (1999), $a = K(1 - m / N)$, where a is the habit strength, $N$is the number of A trials, m is a constant, and K is the asymptote. With the appropriate parameter values (e.g., $m = 0.5$, $K = 10$), there is a large difference in habit strength between 1 and 6 A trials (4.2 units), but a negligible difference between 6 and 11 A trials (0.3 units).

It could be argued that the asymptote is an artifact of random responding after many A trials, perhaps because of fatigue or boredom. Results indicated, however, that infants in all conditions were equally likely to search correctly, and hence systematically, on all B trials following the first correct search at B.

U-Shaped Pattern

A U-shaped pattern of behavior would be consistent with the hierarchical competing systems model (Marcovitch & Zelazo, 1999).[3] On this dual-process model, increases in the number of A trials not only produce increases in habit strength, but also provide opportunities for understanding the task structure (see also Aguiar & Baillargeon, 2000; Munakata, 1998). This latter influence can be understood as a probability, on each trial, of reflecting on the task. Presumably, this probability is low for a 9-month-old infant, but it may be nonzero, and the probability of reflection will be cumulative over trials. More precisely, the probability that reflection occurs once over the entire range of A trials can be formulated as $1 - (1 - R)^N$, where N is the number of A trials and R is the probability on each trial that the infant will reflect on the task structure. If we assign a probability of .1 to R, then the probability of reflection is .1 after 1 A trial, .47 after 6 A trials, and .69 after 11 A trials. Considering the probability of reflection together with habit strength, a U-shaped pattern emerges. After 6 A trials, there is maximal interference from habit strength and less than a 50% chance of reflection. Consequently, the infant will likely persevere. This is in contrast both to the infant who experienced 1 A trial and is not influenced to the same degree by the habit, and to the infant who experienced 11 A trials and has a higher probability of reflection. In both of these cases, one would expect less perseveration than after 6 A trials, providing a U-shaped pattern of behavior.

CONCLUSION

These findings indicate that the number of A trials affects the A-not-B error, and as such, they provide empirical support for a competing systems interpretation of this important developmental phenomenon. From this perspective, the A-not-B paradigm can be used to measure the relative strengths of different kinds of influence, and it may provide a continuous marker of the development of controlled cognitive processes. The further finding that the effect of A trials appears to be asymptotic or U-shaped raises interesting questions concerning the possible role of reflection on task structure. However, future research with additional levels of number of A trials will be required to answer these questions.

Notes:

[1]Strictly speaking, the curve would be U-shaped if the dependent measure were the proportion of infants who searched correctly, and inverted U-shaped if the dependent measure were the proportion of infants who searched incorrectly. However, because this orientation depends on an arbitrary classification of the dependent variable, both types of curves are referred to as U-shaped in this article.

[2]The relatively high attrition rate (42%) can probably be attributed to the fact that our criteria for admission were relatively strict. Whereas some studies do not require correct performance on the A trials to be included in the analyses (e.g., Smith et al., 1999), we required infants to pass both an extensive training phase and the A-trial phase.

[3]In principle, it is possible to explain a U-shaped pattern of behavior using a model where the number of A trials affects only one process (e.g., habit strength). However, we are unaware of any psychologically motivated model that would predict this pattern of behavior.

REFERENCES

Aguiar, A., & Baillargeon, R. (2000). Perseveration and problem solving in infancy. *Advances in Child Development and Behavior*, *27*,135–180.

Bjork, E. L., & Cummings, E. M. (1984). Infant search errors: Stage of concept development or stage of memory development. *Memory and Cognition, 12,* 1–19.

Bremner, J. G., & Bryant, P. E. (1977). Place versus response as the basis of spatial errors made by young infants. *Journal of Experimental Child Psychology*, *23,* 162–171.

Butterworth, G. E. (1977). Object disappearance and error in Piaget's stage IV task. *Journal of Experimental Child Psychology*, *23,* 391–401.

Diamond, A. (1983). *Behavior changes between 6–12 months of age: What can they tell us about how the mind of the infant is changing?* Unpublished doctoral dissertation, Harvard University, Cambridge, MA.

Diamond, A. (200 1). Looking closely at infants' performance, and experimental procedures, in the A-not-B task [Commentary on Thelen et al.]. *Behavioral and Brain Sciences*, *24*,38–41.

Diamond, A., Cruttenden, L., & Neiderman, D. (1994). AB with multiple wells: 1. Why are multiple wells sometimes easier than two wells? 2. Memory or memory + inhibition? *Developmental Psychology*, *30,* 192–205.

Evans, W. F. (1973). *The stage IV error in Piaget's theory of object concept development: An investigation of the role of activity.* Unpublished doctoral dissertation, University of Houston, Houston, TX.

Flavell, J. H., Miller, P. H., & Miller, S. A. (1993). *Cognitive development* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Gratch, G. (1975). Recent studies based on Piaget's view of object concept development. In L. B. Cohen & P. Salapatek (Eds.), *Infant perception from sensation to cognition* (Vol. 2, pp. 51–99). London: Academic.

Harris, P. L. (1986). Bringing order to the A-not-B error. *Monographs of the Society for Research in Child Development*, *51*(214), 52–61.

Harris, P. L. (1987). The development of search. In P. Salapatek & L. B. Cohen (Eds.), *Handbook of infant perception* (Vol. 2, pp. 155–207). New York: Academic.

Harris, P. L. (1989). Object permanence in infancy. In A. Slater & G. Bremner (Eds.), *Infant development* (pp. 103–121). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Hofstadter, M., & Reznick, J. S. (1996). Response modality affects human infant delayed-response performance. *Child Development*, *67*,646–658.

Hull, C. L., Felsinger, J., Gladstone, I., & Yamaguchi, H. (1947). A proposed quantification of habit strength. *Psychological Review*, *54*, 237–254.

Landers, W. F. (1971). Effects of differential experience on infants' performance in a Piagetian Stage IV object-concept task. *Developmental Psychology*, *5,* 48–54.

Marcovitch, S., & Zelazo, P. D. (1999). The A-not-B error: Results from a logistic meta-analysis. *Child Development*, *70*,1297–1313.

Munakata, Y. (1997). Perseverative reaching in infancy: The roles of hidden toys and motor history in the AB task. *Infant Behavior and Development*, *20,* 405–416.

Munakata, Y. (1998). Infant perseveration and implications for object permanence theories: A PDP model of the AB task. *Developmental Science, 1*, 161–184.

Piaget, J. (1954). *The construction of reality in the child.* New York: Basic Books.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Schmuckler, M. A., & Tsang-Tong, H. Y. (2000). The role of visual and body movement information in infant search. *Developmental Psychology*, *36,* 499–510.

Schuberth, R. E. (1983). The infant's search for objects: Alternatives to Piaget's theory of concept development. In L. P. Lipsitt & C. K. Rovee-Collier (Eds.), *Advances in infancy research* (Vol. 2, pp. 137–182). Norwood, NJ: Ablex.

Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: The task dynamics of the A-not-B error. *Psychological Review*, *106,* 235–260.

Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). A dynamic field theory of infant perseverative errors. *Behavioral and Brain Sciences, 24*, 1–86.
Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.
Thorndike, E. L. (1921). *Educational psychology: Briefer course.* New York: Teachers College Press.
Wellman, H. M., Cross, D., & Bartsch, K. (1986). Infant search and object permanence: A meta-analysis of the A-not-B error. *Monographs of the Society for Research in Child Development, 51*(214), 1–51, 62–67.