

Rational higher-order belief revision in young children

Katherine Kimura and Alison Gopnik

University of California, Berkeley

Author Note

Katherine Kimura, Department of Psychology, University of California, Berkeley; Alison Gopnik, Department of Psychology, University of California, Berkeley.

This research was supported by the National Science Foundation Graduate Research Fellowship (DGE-1106400) to KK and by grants from the National Science Foundation (BCS-1331620) and the Bezos Family Foundation to AG. We would like to thank the families who participated in this research as well as the centers at the Early Childhood Education Program at the University of California at Berkeley, the Monteverde School, and the Step One School.

Correspondence concerning this article should be addressed to Katherine Kimura, Department of Psychology, University of California, Berkeley, 2121 Berkeley Way, Berkeley, CA 94720-1650. E-mail: kkimura@berkeley.edu

Word Count: 4,003 (excluding references and title)

Abstract

Belief revision can occur at multiple levels of abstraction, including lower-level and higher-order beliefs. It is unclear, however, how conflicting evidence interacts with prior beliefs to encourage higher-order belief revision. The current study explores how 4- and 5-year-olds ($N = 96$) respond to evidence that directly conflicts with their causal higher-order beliefs. When shown a single event that directly violated a strongly supported prior belief, preschoolers largely maintained their initial higher-order belief. However, when the prior belief was more weakly supported and the counterevidence was stronger, children changed their minds. These findings indicate that young children can revise their higher-order beliefs and, furthermore, do so depending on the strength of both the evidence and their prior beliefs.

Keywords: belief revision, higher-order beliefs, causal learning

Rational higher-order belief revision in young children

Children are remarkable learners. Despite their limited knowledge, children can form higher-order beliefs, or overhypotheses (Goodman, 1955), about the world from surprisingly few examples (e.g., Dewar & Xu, 2010; Kemp, Perfors, & Tenenbaum, 2007; Lucas, Bridgers, Griffiths, & Gopnik, 2014; Sim & Xu, 2017). For instance, children might come to believe that all dogs bark, all cows moo, and all pigs oink, while also forming the higher-order belief that every animal species makes a unique noise that is different from other animal species. This higher-order belief then constrains specific lower-level beliefs. When presented with an unfamiliar animal, children could predict that this animal will produce a sound that is uniquely distinct from barking, mooing, and oinking. Suppose, however, that children observe evidence that violates this prediction (e.g., a barking sea lion). How will they update their higher-order belief to deal with this new information?

Children, including young infants, are sensitive to evidence that directly conflicts with their beliefs. In fact, previous studies suggest that belief violations can enhance learning by encouraging spontaneous explanations (Legare, Gelman, & Wellman, 2010; Legare, Schult, Impola, & Souza, 2016) and selective exploration (Bonawitz, van Schijndel, Friel, & Schulz, 2012; Stahl & Feigenson, 2015; van Schijndel, Visser, van Bers, & Raijmakers, 2015). Even in the absence of explanation and exploration, children can revise their beliefs in response to new evidence. For example, preschoolers can use belief-violating evidence to update their understanding of balance (Bonawitz et al., 2012b), theory of mind (Amsterlaw & Wellman, 2006; Slaughter & Gopnik, 1996), and the logical form of causal relations (Lucas et al., 2014).

Children also appear to normatively revise their beliefs in ways that are consistent with Bayesian learning (Kushnir & Gopnik, 2007). When four-year-olds are asked to activate a novel

toy with a block, they overwhelmingly place the block directly on top of the toy. However, after observing that a block held above the toy was *more* likely to activate the toy than a block placed directly on top, these children revised their prior belief. When asked to activate the toy again, this time with a new block, they appropriately held the object over the toy.

Similarly, Schulz, Bonawitz, and Griffiths (2007) showed that young children will normatively revise their lower-level beliefs about causal relations that cross domains. Preschoolers initially deny that psychological factors (e.g., worry) can result in physical responses (e.g., stomachache), expecting instead that physical outcomes are the result of physical causes (e.g., eating bad food; Notaro, Gelman, & Zimmerman, 2001). If, however, children are shown multiple instances supporting a psychosomatic cause, then four- and five-year-olds will revise their prior belief and endorse this cross-domain causal relation. This learning, however, is relatively conservative, as children will often fail to generalize their revised belief to a new psychosomatic event (e.g., nervousness causes sickness). That is, revising their beliefs about one causal relation does not increase their willingness to accept other similar relations.

Although it is clear that children can revise their beliefs in response to counterevidence, there are at least two aspects of belief revision that remain unclear. First, it is unclear whether children are revising their higher-order beliefs (e.g., psychological factors can cause physical symptoms) or are instead simply updating their lower-level beliefs (e.g., worry can cause stomachaches). If children are revising their higher-order beliefs, then they should be able to apply their updated belief to other lower-level examples provided that these are consistent with and thus supported by their higher-order belief. Alternatively, children might be simply updating their lower-level beliefs, thereby inferring that the observed evidence applies narrowly in scope. Thus, one goal of the current study is to examine whether children can actually update their

higher-order beliefs given conflicting information or whether they will only update their lower-level beliefs.

Second, it is unclear how conflicting evidence interacts with prior beliefs to promote higher-order belief revision. Bayesian learning suggests that the probability that learners will revise their higher-order belief depends on both the prior probability of that belief and the strength of the evidence. If the prior probability is high, then more evidence may be required to overturn the belief. But if the prior probability is low, then small amounts of data should suffice. Notably, however, earlier studies of belief revision examined learners' response to evidence that violates their current theories, such as their beliefs about causal spatial contiguity or psychosomatic causes. These beliefs will vary in strength depending on the individuals' experience, making it unclear whether the decision to revise reflects a normative Bayesian interaction between prior beliefs and observed evidence.

Indeed, children's decisions about whether to revise might be independent of the strength of their prior beliefs and the new evidence. For instance, children might be irrationally inflexible in their belief revision by requiring large amounts of data to change their beliefs (e.g., Zelazo, Frye, & Rapus, 1996). Alternatively, they might be irrationally willing to revise, changing their beliefs regardless of the strength of the evidence. It is also possible, however, that children might show a rational balance between the strength of their prior beliefs and the strength of the new evidence. In the current study, we explore these questions by systematically controlling the evidence children observe for an initial higher-order belief and the evidence that then contradicts that belief.

We investigate whether children can revise their higher-order beliefs and, furthermore, the extent to which children will revise their beliefs by attending to both the weight of their prior

and the causal strength of the evidence. Children were shown pairs of novel toys and were taught a deterministic rule that supported a higher-order belief about how the toys worked. Participants then observed evidence supporting a different deterministic rule that conflicted with the first. Finally, they were asked to activate a novel toy by selecting from among three blocks, which either supported or failed to support either rule. If children revise their higher-order beliefs as a function of the both their priors and the evidence, then they should revise their higher-order beliefs when the conflicting evidence is stronger than the initial evidence and, likewise, maintain their higher-order beliefs when the initial evidence is stronger than the counterevidence. If, however, they are not sensitive to the strength of the evidence, then they should be just as likely to update (or maintain) their higher-order beliefs in both these cases.

Method

Participants

Ninety-six preschoolers ($M = 4;10$, range = 4;0–5;8, 42 females) from the San Francisco Bay Area participated in this study between April 2015 and May 2017, including fifty-seven 4-year-olds ($M = 4;6$, range = 4;0–4;11, 25 females) and thirty-nine 5-year-olds ($M = 5;4$, range = 5;0–5;8, 17 females). Children were randomly assigned to either a condition with one learning trial and four revision trials ($n = 27$ 4-year-olds, 11 females; $n = 23$ 5-year-olds, 9 females) or to a condition with four learning trials and one revision trial ($n = 30$ 4-year-olds, 14 females; $n = 16$ 5-year-olds, 8 females). Children were from predominantly middle- to upper-class families and were recruited from and tested at local preschools. An additional 15 children participated but were excluded from the analysis for experimenter error ($n = 5$), for failure to complete the task ($n = 4$), or for accidental malfunction of the toys ($n = 6$).

Materials

Sixteen toys (6" x 6" x 1.5"), each equipped with a hidden wireless doorbell, were organized into eight identical pairs. Modeled after the blicket detector paradigm (Gopnik & Sobel, 2000) with procedures similar to those used by Sim and Xu (2017), it appeared that certain blocks 'activated' the toys, making them play music. In fact, the toys were surreptitiously activated by the experimenter.

<Figure 1>

Glued to the top of each toy was a flat wooden piece (4.5" x 4.5"). These pieces varied across pairings such that each pair had a unique color and shape that was different from the other pairs. For example, one pair had *blue circles*, whereas another pair had *red squares*. The toys played a different melody depending on the piece on top. For instance, the toys with blue circles played *Happy Birthday* when activated, whereas the toys with red squares played *Twinkle, Twinkle Little Star* (see Figure 1).

Three wooden blocks (1.5" x 1.5" x 1.5") corresponded to each pair of toys: (1) a shape match, (2) a color match, and (3) a distractor. One of the blocks—the shape match—matched the piece on top of the toy in shape but not in color (e.g., a red *circle* block matched the blue *circle* toy), whereas the other block—the color match—matched in color but not in shape (e.g., a *blue* square block matched the *blue* circle toy). The third block—the distractor—did not match the toy in either shape or color (e.g., a green bridge block did not match the blue circle toy in either shape or color). Critically, none of the blocks matched any of the toys on both shape and color.

Design

Participants were randomly assigned to one of two conditions, which differed only in the number of training trials administered during the learning and revision phases. One condition consisted of one training trial in the learning phase and four in the revision phase. In the other

condition, the learning phase had four training trials, whereas the revision phase had only one. The activation rule in each phase (color vs. shape) was counterbalanced across participants. Thus, the experimental design was 2 (training trials: 1 vs. 4) x 2 (rule: color vs. shape) with the number of training trials in each phase and the activation rule as between-subjects variables.

Procedure

Children were seated at a table directly across from the experimenter. Each session lasted approximately 15 minutes and consisted of two phases: (1) a learning phase and (2) a revision phase. In the learning phase, children saw evidence that supported a particular higher-order belief about how the toys worked. This established a common set of priors for all the children. In the revision phase, they were shown evidence supporting a different higher-order belief that conflicted with the previous one. For example, children might initially learn that the toys activate when the blocks match the toys' top piece in shape but not in color; however, in the revision phase, children would observe the toys activate with blocks matching in color rather than shape.

Across the two phases, children selected seven pairs of toys: five were observed as training trials, whereas two were used for test trials. Once a pair of toys was selected and the trial was completed, that pairing was removed from the selection for subsequent trials so that a new pair of toys was used for each trial. Both the learning and revision phases began with either one or four training trials and ended with a single test trial.

Learning phase. The experimenter began the learning phase by introducing children to her toys: "Look, these are my toys! Sometimes when I put blocks on top of my toys they play music. And sometimes when I put blocks on top of my toys they do not play music." The experimenter then invited the child to play her game and figure out how the toys worked.

For each training trial of the learning phase, the experimenter presented photographs, one for each pair of toys, and asked the child to select a pair by pointing to its picture. By asking children to select the toys and, hence, decide the presentation order, we discouraged the inference that the experimenter had already separated toys that worked by one rule from those that worked by another. After the child responded, the experimenter placed the selected toys on the table and placed two corresponding blocks—the shape match and the color match—directly in front of the first toy, randomizing the left-right placement across trials. The experimenter then, one at a time, placed each block on top of the toy, which activated according to either a shape rule (e.g., *circle* blocks activate *circle* toys) or a color rule (e.g., *blue* blocks activate *blue* toys). To emphasize the outcome, the experimenter would smile and say, “It turned on! It’s playing music!” when the toy activated, whereas she would shake her head and say, “No music. It didn’t turn on.” when it did not activate. After the effect of both blocks was demonstrated—one which resulted in activation and one that did not—the experimenter moved the first toy aside and placed the same blocks on the second toy, again doing so one at a time. This toy, which was identical to the first one, activated according to the same rule. The experimenter then removed these toys and, for children who observed four training trials, repeated the procedure with three new pairs of toys.

After completing the training trials, the experimenter administered a forced-choice test trial to assess whether or not the child learned the causal rule and, critically, generalized this rule to other novel toys. Similar to the training trials, the test trial began with the experimenter asking the child to select a new pair of toys by pointing to its picture. However, unlike the training trials, the experimenter presented only one of the two toys along with three test blocks—the shape match, the color match, and the distractor. These blocks were arranged on a tray in a

random presentation order in front of the toy. The experimenter then asked the child to demonstrate his or her higher-order belief about the toys by asking, “Can you point to what will make my toy play music?” as she moved the tray towards the child. After the child responded, the experimenter provided a neutral response, removed the toy and blocks, and proceeded to the revision phase.

Revision phase. The revision phase was identical to the learning phase with the following exception: the training trials in the revision phase supported the opposite rule to that demonstrated in the learning phase. For instance, if children initially learned that blocks of the same color activated a set of toys, then in the revision phase they observed that blocks of the same shape activated the remaining set of toys, while blocks of the same color did not. As in the learning phase, a final forced-choice test trial assessed whether participants revised or maintained their original higher-order belief.

Coding

Responses were coded based on whether the child selected the block corresponding with the observed rule for that given phase. For example, children who observed the color rule and selected the color match in the learning phase were coded as learning the higher-order rule, whereas children who selected the shape match or the distractor after observing identical evidence were coded as not. In the revision phase, responses were similarly coded as being consistent or inconsistent with the revision rule. The first author performed the coding, whereas a second coder blind to the condition and activation rule scored a random subset of participants (40%). Interrater reliability was 100%.

Results

Preliminary analyses revealed no significant difference between children who observed the color rule and selected the color match from those who observed the shape rule and selected the shape match, thus implying that children were just as likely to learn the color rule as they were the shape rule. There were also no significant differences in performance between four- and five-year-olds. The data were, therefore, collapsed across learning rules and across age groups. All analyses were two-tailed.

Learning phase. Despite this markedly challenging task, especially for participants who observed a single training trial, children selected the block corresponding with the learning rule more often than would be expected by chance (i.e., .33) after observing one training trial (30 out of 50 children; 60.0%) $p < .001$, binomial test), or four training trials (32 out of 46 children; 69.6%), $p < .001$, binomial test. Although children who observed four training trials observed more evidence in support of the learning rule, the difference between the two conditions was not significant, $p = .395$, Fisher's exact test.

Although this consistent and correct responding suggests that children inferred the appropriate causal relation, learning a higher-order rule from few observations was, nevertheless, a challenging task. When they saw only one training trial, 20 children out of 50 (40.0%) failed to infer the correct rule and, even when they saw four consecutive training trials that supported the same higher-order rule, 14 children out of 46 (30.4%) still failed to learn the rule.

Revision phase. We next examined whether or not children revised their causal beliefs given the strength of the conflicting evidence. To do this, we restricted analyses to include only those children who learned the initial rule (one learning trial condition: $n = 30$; four learning trials condition: $n = 32$) and tested whether they varied their revision patterns as a function of the number of training trials (i.e., 1 vs. 4 trials). If participants were weighing the strength of the

evidence against the strength of their prior beliefs, then they should be *more* likely to revise their beliefs when shown one learning and four revision trials than if shown four learning and one revision trial.

Consistent with these predictions, we found that children who observed one learning trial and four revision trials were more likely to select the block that corresponded with the revision rule (19 out of 30 children; 63.3%) than those who observed four learning trials and one revision trial (7 out of 32 children; 21.9%), $p = .002$, Fisher's exact test. Furthermore, when presented with one learning and four revision trials, children selected the block corresponding with the revision rule at a rate that exceeded chance performance, $p < .001$, binomial test, indicating that they reliably revised their higher-order beliefs. In contrast, children who observed four learning trials and one revision trial were more likely to select the block that was consistent with the learning rule (22 out of 32 children; 68.8%), $p < .001$, binomial test. In other words, when the prior was weaker and the counterevidence was stronger, children were likely to revise their beliefs appropriately. However, when the prior was stronger and the counterevidence was weaker, children largely maintained their initial higher-order beliefs.

Discussion

These analyses indicate that children can revise their higher-order beliefs in a way that is consistent with Bayesian learning. When children observed pairs of novel toys that activated according to a particular rule, thereby supporting a higher-order belief, they were more likely to maintain this belief if, overall, they viewed more evidence that supported their initial rule than the revised rule. However, if they instead observed more evidence that supported the conflicting rule, then children would systematically revise their belief to appropriately account for the new data.

Similar to previous studies (e.g., Bonawitz et al., 2012b; Kushnir & Gopnik, 2007; Legare et al., 2016; Lucas et al., 2014; Schulz et al., 2007), we found that children can quickly revise their beliefs in response to conflicting evidence. However, unlike previous studies, we explicitly demonstrate that children can revise their higher-order beliefs. For example, observing the blue block activate the blue toys prompts children to form a higher-order belief—e.g., blocks that correspond in color will activate the toys. This higher-order belief then constrains and guides their beliefs about novel toys in the test trial. For instance, children will predict that the red block will activate the red square toy rather than the square block. If, however, they observe that the red block fails, whereas the square block succeeds—and they also observe other instances that support the alternative shape rule—then children will update their higher-order belief. On the test trial, they will now predict that blocks corresponding in shape, and not in color, will activate the toys.

Limitations and Future Directions

Although our findings suggest that children revised their higher-order beliefs, we do not know whether they were also revising their lower-level beliefs. Children might, for instance, continue to rightfully believe that blue blocks activate blue circle toys even though this lower-level belief conflicts with their newly revised higher-order belief. It is also possible that children are updating both their lower- and higher-order beliefs. In addition to now believing that blocks matching in shape activate the toys, children might, for instance, believe that circle blocks activate blue circle toys. In fact, lower-level belief revision might precede higher-order belief revision, particularly if higher-order beliefs are more stable and less susceptible to change. Alternatively, higher-order beliefs, which often have few competing alternatives (Perfors, Tenenbaum, Griffiths, & Xu, 2011), could be revised first. These questions are important in

deepening our understanding about how observed evidence interacts with complementary and competing beliefs but answering them will depend on further research.

We have focused on how learners respond to evidence that behaves deterministically; however, the evidence we typically encounter behaves stochastically. How might learners reason about such probabilistic evidence when revising their higher-order beliefs? Our experiment does not directly address this question, but it seems plausible that the strategies learners employ when they evaluate deterministic evidence are similar to those they use when they weigh probabilistic evidence. Indeed, children might have reasoned that the toys behaved probabilistically when they combined the evidence from both the learning and revision phases. In fact, children might think that the system instantiated a higher-order regularity and that there were occasional exceptions to that regularity—i.e., the one conflicting trial might be a fluke (Nosofsky, Palmeri, & McKinley, 1994). As a result, children would maintain their initial belief when the evidence strongly favored that belief—even when there is an occasional exception.

Although we found no difference between four- and five-year-olds in their ability to revise their higher-order beliefs, many children failed to initially learn the higher-order belief and, therefore, could not be included in the revision phase. It is thus possible that the children we tested—i.e., those who learned the higher-order rule—were particularly advanced compared to four- and five-year-olds in general. There is also evidence suggesting that younger children, three-year-olds in particular, might have more difficulty revising their beliefs in response to new data. In Kushnir & Gopnik (2007), three-year-olds, like four-year-olds, initially believed that causation required direct contact. Although the older children eventually revised this belief when they saw counterevidence, the younger children had difficulty overriding their prior preference, even after they saw deterministic, unambiguous evidence to the contrary. Similarly, three-year-

olds, in contrast to older children, failed to endorse psychosomatic causes and instead continued to demonstrate a strong bias towards the domain-appropriate cause. Younger children only revised their beliefs after a two-week training, during which they were taught to reason about statistical evidence (Bonawitz, Fisher, & Schulz, 2012). This is consistent with other evidence suggesting that three-year-olds are relatively inflexible. For instance, when preschoolers are instructed to sort a series of bi-dimensional cards first according to one dimension and then according to another, they often perseverate and fail to switch to the second dimension (Zelazo et al., 1996). In future studies, we plan to extend this work to younger children to explore whether there is a qualitative difference between three- and four-year-olds, as suggested by the earlier studies, and whether this difference is related to performance on dimensional shift tasks. It would also be interesting to explore other potential correlates of this kind of flexibility, such as bilingualism, culture, socioeconomic status, or executive control.

Conclusion

Belief revision is a highly complex and dynamic process that involves evaluating the evidence against prior beliefs at multiple levels of abstraction. Despite these challenges, our findings suggest that children can revise their higher-level beliefs when they see conflicting evidence and, indeed, do so rationally. When the evidence strongly supports an alternative, children overwhelmingly revise their beliefs to reflect the new information. These higher-order beliefs, however, are relatively robust. When there is more evidence for the initial belief than for the alternative, children maintain their initial belief. Although our knowledge is relatively stable, every once in a while, when counterevidence evidence is strong enough, it can lead to a new discovery that radically alters our thinking.

References

- Amsterlaw, J., & Wellman, H. M. (2006). Theories of mind in transition: A microgenetic study of the development of false belief understanding. *Journal of Cognition and Development, 7*, 139-172. https://doi.org/10.1207/s15327647jcd0702_1
- Bonawitz, E., Fisher, A., & Schulz, L. (2012) Teaching three-and-a-half-year-olds to revise their beliefs given ambiguous evidence. *Journal of Cognition and Development, 13*, 266-280. <https://doi.org/10.1080/15248372.2011.577701>
- Bonawitz, E. B., van Schijndel, T. J., Friel, D., & Schulz, L. (2012). Children balance theories and evidence in exploration, explanation, and learning. *Cognitive Psychology, 64*, 215-234. <https://doi.org/10.1016/j.cogpsych.2011.12.002>
- Dewar, K. M., & Xu, F. (2010). Induction, overhypothesis, and the origin of abstract knowledge: Evidence from 9-month-old infants. *Psychological Science, 21*, 1871-1877. <https://doi.org/10.1177/0956797610388810>
- Goodman, N. (1955). *Fact, fiction, and forecast*. Cambridge, MA: Harvard University Press.
- Gopnik, A., & Sobel, D. M. (2000). Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child Development, 71*, 1205–1222. <https://doi.org/10.1111/1467-8624.00224>
- Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science, 10*, 307-321. <https://doi.org/10.1111/j.1467-7687.2007.00585.x>
- Kushnir, T., & Gopnik, A. (2007). Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial

- assumptions. *Developmental Psychology*, *43*, 186-196. <http://dx.doi.org/10.1037/0012-1649.43.1.186>
- Legare, C. H., Gelman, S. A., & Wellman, H. M. (2010). Inconsistency with prior knowledge triggers children's causal explanatory reasoning. *Child Development*, *81*, 929-944. <https://doi.org/10.1111/j.1467-8624.2010.01443.x>
- Legare, C. H., Schult, C. A., Impola, M., & Souza, A. L. (2016). Young children revise explanations in response to new evidence. *Cognitive Development*, *39*, 45-56. <https://doi.org/10.1016/j.cogdev.2016.03.003>
- Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition*, *131*, 284-299. <https://doi.org/10.1016/j.cognition.2013.12.010>
- Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*, 53-79. <http://dx.doi.org/10.1037/0033-295X.101.1.53>
- Notaro, P. C., Gelman, S. A., & Zimmerman, M. A. (2001). Children's understanding of psychogenic bodily reactions. *Child Development*, *72*, 444-459. <https://doi.org/10.1111/1467-8624.00289>
- Perfors, A., Tenenbaum, J. B., Griffiths, T. L., & Xu, F. (2011). A tutorial introduction to Bayesian models of cognitive development. *Cognition*, *120*, 302-321. <https://doi.org/10.1016/j.cognition.2010.11.015>

Schulz, L. E., Bonawitz, E. B., & Griffiths, T. L. (2007). Can being scared cause tummy aches?

Naive theories, ambiguous evidence, and preschoolers' causal inferences. *Developmental Psychology*, *43*, 1124-1139. <http://dx.doi.org/10.1037/0012-1649.43.5.1124>

Sim, Z. L., & Xu, F. (2017). Learning higher-order generalizations through free play: Evidence from 2-and 3-year-old children. *Developmental Psychology*, *53*, 642-651.

<http://dx.doi.org/10.1037/dev0000278>

Slaughter, V., & Gopnik, A. (1996). Conceptual coherence in the child's theory of mind:

Training children to understand belief. *Child Development*, *67*, 2967-88.

doi:10.2307/1131762

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, *348*, 91-94. doi: 10.1126/science.aaa3799

van Schijndel, T. J., Visser, I., van Bers, B. M., & Raijmakers, M. E. (2015). Preschoolers perform more informative experiments after observing theory-violating evidence. *Journal of Experimental Child Psychology*, *131*, 104-119.

<https://doi.org/10.1016/j.jecp.2014.11.008>

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, *11*, 37-63. [https://doi.org/10.1016/S0885-](https://doi.org/10.1016/S0885-2014(96)90027-1)

[2014\(96\)90027-1](https://doi.org/10.1016/S0885-2014(96)90027-1)

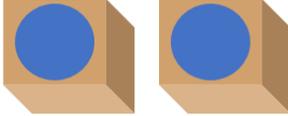
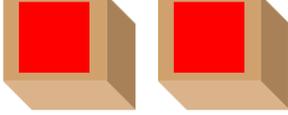
Toys	Shape Match	Color Match	Distractor	Melody
				Happy Birthday
				Twinkle, Twinkle Little Star
				Yankee Doodle

Figure 1. Sample stimulus sets.