What else could happen? Two-, three-, and four-year-olds use variability information to infer novel causal outcomes

Mariel K. Goddu (goddu@berkeley.edu)
Department of Psychology, 2121 Berkeley Way
Berkeley, CA 94720 USA

Trisha Katz (tkatz@ucsd.edu)
Department of Psychology, Department of Psychology, McGill Hall #0109
9500 Gilman Drive, La Jolla, CA 92093 USA

Caren M. Walker (carenwalker@ucsd.edu)
Department of Psychology, Department of Psychology, McGill Hall #0109
9500 Gilman Drive, La Jolla, CA 92093 USA

Abstract

Young children rapidly infer causal relations by tracking contingencies between causes and their effects, and can generalize these rules to novel instances of the same cause. However, this is distinct from the ability to make inferences about whether a particular cause is likely to produce novel effects. Here, we investigate the development of two-, three-, and four-year-olds' ability to recognize and use information about a cause's *variability* to make predictions about other novel outcomes it might produce. Experiment 1 finds that children as young as two years of age infer that a cause that has produced *variable*, rather than *deterministic* outcomes is more likely to produce a novel, previously unobserved effect. Experiment 2 finds that four-year-olds, but not two- and three-year-olds, infer that a *higher variability* cause is more likely to produce a novel outcome than a *lower variability cause.*

Keywords: cognitive development; causal reasoning; inference; probability; variability; causal intervention

Introduction

Learners frequently make predictions about events they have never observed. One way they do this is by learning and generalizing the rules that govern a particular system of causes and effects: e.g., if a red triangle, red circle, and red square block all cause a machine to play music, but a blue triangle block does not, inferring the causal rule "red makes it go" can serve as a guiding principle for determining which other events are likely to occur. For example, the learner may predict that a novel red rectangle will also cause music, but that a blue rectangle will not. This ability to infer causal rules from evidence is well documented in the literature on early causal learning (see Gopnik, 2012; Gopnik & Wellman, 2012 for reviews).

However, not every situation affords the opportunity to rely on concrete causal rules to support future inference. Instead, the outcomes of a single causal intervention may vary across instances; that is, the same action may produce different effects on different occasions, even within the same context. For example, perhaps exclaiming, "LOOK!" on the playground has reliably caused my parent to attend to me. In this case, I may have little reason to suspect that performing this action will result in other possible effects. But, perhaps I discover that performing the same action produces different outcomes at different times: perhaps "LOOK!" sometimes causes my parent to look, sometimes causes other parents to look, and sometimes causes my younger sibling to come over to see what I am doing. This set of variable outcomes not only licenses my expectation that those same, particular outcomes may recur, but also that other, *new* things may happen—e.g., that yelling "LOOK!" might cause other children on the playground to crowd around me. This inference is not governed by learning and applying a particular causal rule, but rather by appreciating and forming an abstract generalization about the variable effects that this single action can generate.

Here, we investigate the possibility that even very young children are able to learn abstract information about the *variability* of a cause and use it to inform their understanding of what other events might be possible. To do this, children must be able to track and use probabilistic data to infer causal rules, distinguish between variable and deterministic evidence, use variability information in the environment to inform their inferences, and generate abstract hypotheses about the types of causes that are likely to vary. Below, we briefly review prior research suggesting that children in fact possess these requisite skills. We then introduce the current experiments, which integrate these previous findings to ask whether young children are able to form an abstract hypothesis about causal variability to support inferences about previously unobserved outcomes. Specifically, will children infer that a cause that has produced *variable* outcomes is more likely to produce a *novel* outcome than a cause that does not vary (Experiment 1), or varies to a lesser extent (Experiment 2)?

Variability information shapes children's inferences

The literature on the development of causal reasoning in early childhood demonstrates that children as young as 16 months can learn from patterns of statistical contingency to infer causal rules, intervene to produce desired effects, and design novel interventions (e.g., Gopnik et al., 2001, 2004; Gopnik & Sobel, 2000; Gweon & Schulz, 2011; Kushnir & Gopnik, 2007; see Gopnik, 2012 and Gopnik & Wellman, 2012 for reviews). In the majority of studies, participants observe a sequence of evidence that supports a deterministic causal rule—e.g., "red blocks, but not blue blocks, cause a machine to play music." However, children can also learn from probabilistic evidence: If a red block causes a machine to play music 66% of the time, and a blue block causes the machine to play music 33% of the time, four-year-olds will choose the "stronger" block when given the opportunity to produce the causal outcome themselves (Kushnir & Gopnik, 2005). Other evidence suggests that children as young as 24-months can discriminate between probabilistic and deterministic causes (Waismayer & Meltzoff, 2017).

However, the ability to track statistical contingencies to infer the relative likelihood and strength of a cause is distinct from the ability to make inferences based on a cause's *variability*—the extent to which it produces novel outcomes. There are several prior studies that suggest that children may be sensitive to variability information. For example, Ahl and Keil (2017) find that four- to eight-year-olds infer that a machine that performs more variable functions (e.g., making muffins and soups) will have more internal complexity than a machine that performs the same number of less variable functions (e.g., making muffins and cupcakes) (2017; Ahl et al., 2018; see also Erb, Buchanan, & Sobel, 2013).

Other evidence demonstrates that children are also sensitive to variability information in the context of decision-making. Specifically, recent work investigating preschoolers' behavior in a modified multi-armed bandit task demonstrates that they prefer to explore the variable outcomes afforded by taking several different actions over "exploiting" those actions that produce the greatest material gains (Sumner, Streyvers, & Sarnecka, 2019; Sumner et al., 2019). Similarly, four-year-olds appear to be sensitive to variability in the behavior (i.e., trustworthiness) of social agents, and use this information when deciding whether to "cash out" on an immediate reward or wait to maximize it (Kidd, Palmeri, & Aslin, 2013).

Finally, a large body of research demonstrates that children as young as eight months are sensitive to the relation between the variability of samples and populations. For example, infants are surprised when a sample of mostly white balls is drawn from a population of mostly red balls, and they can compare probabilistic distributions to maximize their chances of accessing a desired reward (Xu & Garcia, 2008; Xu & Denison, 2009).

Children learn abstract principles that guide future inferences

Prior research demonstrates that even very young children can learn general, abstract principles that lead them to privilege particular *types* of hypotheses over others (e.g., Goodman, Ullman, & Tenenbaum, 2011). This mechanism enables learners to make inferences that go beyond their direct experience. For example, 10-month-old infants who observe red shapes emerge from one box, green shapes emerge from another, and yellow shapes emerge from a third, will be surprised if shapes drawn from a fourth box do not conform to a uniform distribution based on color (Dewar & Xu, 2010). Similarly, young children infer that certain types of object properties (e.g., shape) are more relevant than others for determining the extensions of words and categories (Kemp, Perfors, & Tenenbaum, 2007).

In addition to learning these abstract rules, or "overhypotheses," that constrain and guide inference in a variety of non-causal domains, children are also able to infer second-order generalizations about the abstract "form" of a cause—i.e., the *kind* of cause that is most likely to produce a particular outcome. For example, four- to six-year-olds learn to make abstract causal attributions about whether an agent's traits (e.g., their disposition) is more likely than some feature of the environment (e.g., a safety hazard) to explain a set of observations (e.g., a character's reluctance to play on a trampoline) (Seiver, Goodman, & Gopnik, 2013). Similarly, four- to six-year-olds are able to flexibly infer conjunctive causation (e.g., that two blocks are required to make a machine play music) or disjunctive causation (e.g., that only one block is required), depending on the evidence they observe (Lucas et al., 2015). Finally, children as young as 18 months are able to infer higher-order *relational* causal rules from evidence they observe (e.g., that "same" or "different" blocks cause a machine to play music) (Carstensen et al., 2019; Walker, Bridgers, & Gopnik, 2016; Walker & Gopnik 2014, 2017), and in the context of their own free play (Sim & Xu, 2017).

The present experiments

In the present experiments, we go beyond this prior work to investigate whether children are able to infer a more general rule or hypothesis about a different type of abstract feature: the relative *variability* of two causes. If a learner interprets the occurrence of different outcomes (produced by a single cause) as evidence for variability, then this may license the inference that this cause is also likely to generate a previously unobserved outcome. *Experiment 1* investigates whether children as young as two years are able to discriminate between causes that generate deterministic versus variable outcomes and infer that the variable outcome is more likely to produce a novel effect. *Experiment 2* investigates whether children infer that a *high* variability cause is more likely to produce a novel outcome than a *low* variability cause.

Experiment 1

Experiment 1 investigated whether two-, three-, and four-year-olds would infer that a *variable* cause was more likely to produce a novel effect than a *deterministic* cause. If children exclusively attend to the concrete outcomes produced by the two causes, then they should be equally likely to select either the deterministic or variable cause when prompted to produce a *novel* outcome. If, however, children form a more general hypothesis about each cause's variability, then they should correctly infer that the variable cause is more likely to produce an outcome that they have not yet observed.

Methods

Participants and Design A total of 87 participants, including 29 two-year-olds (M_{age} = 30.7 months, *SD* = 3.3), 29 three-year-olds (M_{age} = 41.9 months, *SD* = 2.8), and 29 four-year-olds (M_{age} = 56. months, *SD* = 3.7) were recruited from children's museums and preschools. Fourteen additional children were tested, but excluded due to experimental error (8), issues with language comprehension (2), failure to respond (2), interference by a parent or sibling (1), or inattention (1).

Stimuli and Procedure Participants were introduced to a puppet named Pete the Parrot. The experimenter explained that Pete enjoyed eating yellow and blue pebbles. Children were then provided with the opportunity to feed Pete one pebble of each color, and Pete expressed his satisfaction, saying, "*nom, nom, nom!*".

Next, the experimenter introduced the "birdfeeder," an apparatus composed of four vertical, transparent tubes that each emptied into a transparent jar. The openings at the top of the tubes were occluded from the participants' view, but were accessible to the experimenter. The experimenter told each participant that Pete the Parrot "likes to come to the birdfeeder to eat," and noted that that there were "*one*, *two*, *three*, *four* tubes that the pebbles can fall out of" (pointing to each tube in turn). Children were then introduced to two large buttons of different colors (blue and yellow) that controlled the feeder, each attached to the feeder with wire (see *Figure 1* for a schematic of the stimuli and procedure).

The experimenter introduced each button, one at a time, saying, "I have this [*blue/yellow*] button, and when you press this button, it's going to make the pebbles fall out of the tubes for Pete to eat!" The experimenter prompted the child to press the button, saying, "Go ahead and hit the [*blue/yellow*] button. Let's see what happens!" When the participant hit the button, the experimenter surreptitiously placed a blue pebble into the top of one of the tubes, such that it dropped down into the jar beneath it. The experimenter said, "Wow, cool! Can you press the [*blue/yellow*] button again?" The child was prompted to press the button two mores time, so a total of three pebbles matching the color of the button were released from the feeder. The experimenter said, "Cool! So that's what happened when we pressed the [*blue/yellow*] button. Now let's find out what happens when we press the [*yellow/blue*] button!" The experimenter repeated the same procedure with the second button.

Of the two buttons, one was *deterministic* and the other was *variable.* The *deterministic button* caused color-matched pebbles to fall from the same tube three times in a row (tube 1-1-1), producing one pebble with each press. The *variable button* caused color-matched pebbles to fall in one of two pseudo-random orders (tube 1-4-3, tube 4-3-1), counterbalanced between subjects. Critically, however, *neither button ever caused a pebble to drop from tube 2.* Because the jars were clear, participants were always able to see the complete distribution of blue and yellow pebbles. The right-left placement of the buttons, the order of presentation (deterministic or variable button first), and the color of button/pebbles corresponding to deterministic or variable causation were all counterbalanced across subjects.

At test, the experimenter retrieved and animated the Pete the Parrot puppet. The experimenter said, "I think Pete the Parrot is hungry, let's find out which tube Pete wants to eat from!" Pete then "flew" to tube 2. The experimenter said, "Oh! It looks like Pete the Parrot wants to eat from this tube right here. So, he wants the pebbles to come out of this tube right here." The experimenter then presented the two buttons once more, and asked, "Which button should we press to get the pebbles to come out of that tube?" The participants then had the opportunity to press one of the buttons, and their choices were recorded. If a participant said they wanted to press both buttons, they were prompted to choose just one.

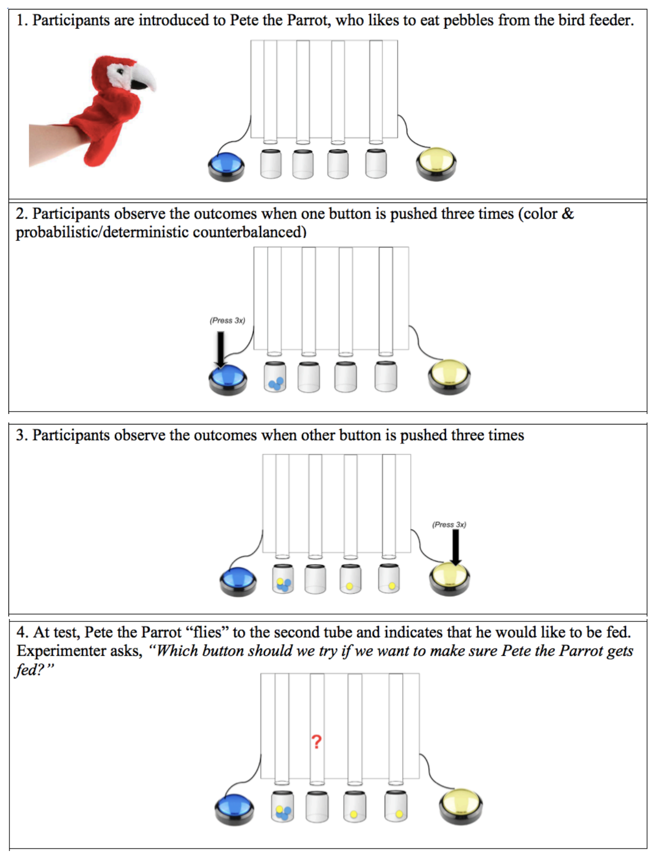

Figure 1: Schematic of Experiment 1 stimuli and procedure

Results and Discussion

The results of Experiment 1 suggest that participants of all age groups distinguished between the variable and deterministic causes and used this information to infer which would produce the desired novel outcome. A significant majority of children (73.6%) chose to intervene on the variable button to produce the novel causal outcome, X^2 (1, 87) = 19.32, p < .001, with no significant difference among two-year-olds (69%), three-year-olds (76%), and four-year-olds (76%), X^2 (2, 87) = 0.47, p = .79.

Children's selection of the variable cause may have reflected their early understanding that variable causes are more likely to produce novel outcomes. However, an alternative interpretation does not require this inference. Instead, children may have inferred that deterministic causes are unlikely to produce novel outcomes, and simply avoided the deterministic button. That is, children may have chosen the variable button, either because they thought it was more likely to produce a new outcome, *or* because they recognized that the deterministic button was a "bad bet."

Experiment 2 thus replaced the *deterministic* cause with a *low-variability* cause to investigate whether children would infer that greater variability indicates a higher likelihood of generating a novel outcome.

Experiment 2

Experiment 2 investigated whether two-, three-, and four-year-olds' would infer that a *higher variability* cause was more likely to produce a novel effect than a *lower variability* cause.

Methods

Participants and Design A total of 87 participants, including 29 two-year-olds (M_{age} = 29.3 months, SD = 3.0), 29 three-year-olds (M_{age} = 40.1 months, SD = 3.1), and 29 four-year-olds (M_{age} = 52.6 months, SD = 3.3) were recruited from children's museums and preschools. Eight additional children were tested, but excluded due to experimental error (5) or failure to respond (3).

Stimuli and Procedure The stimuli were identical to that of Experiment 1, and the procedure was highly similar. However, there were two important differences. First, the number of button-presses for each type of cause was increased to five to maximize the observable difference in variability between the high- and low-variability causes. Second, the deterministic button from Experiment 1 was replaced by a *low-variability* button, which produced pebble drops in the pseudorandom orders 1-3-1-1-1 and 1-1-1-3-1 (counterbalanced across participants). The pattern of pebble-drops for the high-variability cause alternated between three pseudorandom orders: 1-3-1-4-3, 4-1-3-4-1, and 3-4-1-4-3. As in Experiment 1, participants never observed either button produce a pebble from tube 2. At test, participants were again asked which button to press in order to make a pebble fall out of Pete's preferred tube (tube 2) to feed him.

Results and Discussion

The results of Experiment 2 indicate that four-year-olds correctly discriminated between low- and high-variability causes, and inferred that the high-variability cause was more likely to produce a novel, previously unobserved outcome: 72.4%, selected the high-variability button, X^2 (1, 29) = 5.83, p = .02, which was not significantly different from their performance in Experiment 1 (76%), X^2 (2, 58) = .09, p = .76. By contrast, two- and three-year-olds performed at chance: 58.6% of two-year-olds, X^2 (1, 29) = 0.86, p = .35, and 51.7% of three-year-olds, X^2 (1, 29) = 0.03, p = .86 selected the high-variability button (see *Figure 2*).

General Discussion

Taken together, the findings of Experiments 1 and 2 suggest a developmental shift in children's ability to use information about a cause's variability as a guide to future outcomes. *Experiment 1* demonstrates that two-, three-, and four-year-olds are sensitive to the difference between a variable and deterministic cause, and infer that a variable cause is more likely to produce a novel, previously unobserved outcome than a deterministic cause. *Experiment 2* suggests that four-year-olds, but not two- and three-year-olds, discriminate between a high- and low- variability cause to infer that the cause with *greater* variability is more likely to generate something new.

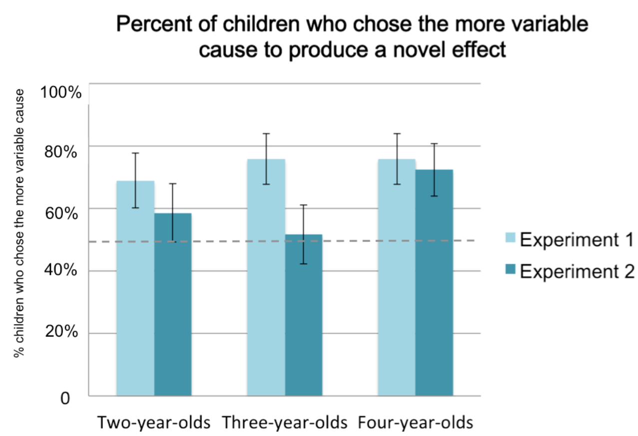

Figure 2: Results of Experiments 1 and 2. Error bars are standard errors.

Given the causal reasoning literature (reviewed above), which demonstrates toddlers' and preschoolers' ability to learn causal rules from both probabilistic and deterministic evidence, it seems improbable that the younger children's performance in Experiment 2 was due to a failure to discriminate the variability of the two causes; more likely, they did not interpret higher variability as a strong enough cue to infer a novel outcome over lower variability. It is also possible that their success in Experiment 1 was not due to drawing an inference about variability, but rather an inference about determinism. In other words, younger children may have (correctly) inferred that a deterministic cause is unlikely to produce a novel outcome, leading them to avoid it and select the only other option available.

Additionally, given that the total number of causal outcomes increased from 6 to 10 between Experiments 1 and 2, it is also possible that younger children simply had more difficulty tracking the evidence. Thus, ongoing work aims to replicate Experiment 1 using 10 trials (i.e., 5 demonstrations each of the deterministic and variable buttons) to match the number of trials in Experiment 2.

These findings also complement a recent theoretical proposal of causal reasoning that highlights young children's "search for invariance" (Lapidow & Walker, 2019). According to this account, children are motivated to discover evidence for invariance—the extent to which a particular causal relation continues to hold across contexts—since it supports broad generalization. Here, we explored whether children are sensitive to information about causal (in)variance within a *single* context. The early appearance of this ability suggests that discriminating deterministic from variable causes may also be critical for reasoning about novel possibilities within a particular causal system.

In sum, while much of the previous work in early causal reasoning focuses on children's ability to infer causal rules from evidence, the present findings bear on a complementary and previously unexamined aspect of early causal inference. That is, in the absence of predictive causal rules or prior knowledge about a potential mechanism, we can rely on the concept of variability to guide our inferences about the types of causes that are likely to produce novel outcomes. Our findings provide evidence that this ability may be in place as early as two years of age. This skill may contribute to children's developing capacity to consider what types of events might happen as they solve problems and imagine new possibilities.

Acknowledgments

Thank you to Jess Wallach for her assistance with data collection. We gratefully acknowledge Birch Aquarium, Fleet Science Center, and the New Children's Museum for allowing data collection for these projects in their fine facilities. We also thank the teachers and staff at Honey Bear Preschool, Maple Hill Early Learning Center, and Breezy Knoll Childcare Center for hosting data collection. Finally, we thank the participants' parents for allowing their children to take part in this research, and the participants themselves for sharing their time and attention.

References

Ahl, R., DeAngelis, E., Stephenson, A., Joo, S., & Keil, F. (2018). It's Complicated: Children Identify Relevant Information About Causal Complexity. *Proceedings of the cognitive sciences society 2018*, *40.*

Ahl, R. E., & Keil, F. C. (2017). Diverse effects, complex causes: children use information about Machines' functional diversity to infer internal complexity. *Child development*, *88*(3), 828-845.

Carstensen, A., Zhang, J., Heyman, G. D., Fu, G., Lee, K., & Walker, C. M. (2019). Context shapes early diversity in abstract thought. *Proceedings of the National Academy of Sciences*, *116*(28), 13891-13896.

Denison, S., Reed, C., & Xu, F. (2013). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5-and 6-month-olds. *Developmental Psychology*, *49*(2), 243.

Denison, S., & Xu, F. (2010). Twelve to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science, 13*(5), 798–803.

Dewar, K. M., & Xu, F. (2010). Induction, overhypothesis, and the origin of abstract knowledge: Evidence from 9-month-old infants. *Psychological Science*, *21*(12), 1871-1877.

Erb, C. D., Buchanan, D. W., & Sobel, D. M. (2013). Children's developing understanding of the relation between variable causal efficacy and mechanistic com- plexity. Cognition, 129, 494–500. doi:10.1016/j.cognition. 2013.08.00

Goodman, N. D., Ullman, T. D., & Tenenbaum, J. B. (2011). Learning a theory of causality. *Psychological review*, *118*(1), 110.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological review*, *111*(1), 3.

Gopnik, A., & Sobel, D. M. (2000). Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child development*, *71*(5), 1205-1222.

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental psychology*, *37*(5), 620.

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological bulletin*, *138*(6), 1085.

Gweon, H., & Schulz, L. (2011). 16-month-olds rationally infer causes of failed actions. *Science*, *332*(6037), 1524-1524.

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental science*, *10*(3), 307-321.

Kidd, C., Palmeri, H., & Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition*, *126*(1), 109-114.

Kushnir, T., & Gopnik, A. (2005). Young children infer causal strength from probabilities and interventions. *Psychological science*, *16*(9), 678-683.

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*(1), 186.

Lapidow, E., & Walker, C. M. The Search for Invariance: Repeated Positive Testing Serves the Goals of Causal Learning.

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*(2), 284-299.

Rhodes, M., Gelman, S. A., & Brickman, D. (2010). Children's attention to sample composition in learning, teaching and discovery. Developmental Science, 13, 421–429. doi:10.1111/j.1467-7687.2009.00896.x

Schulz, L. E., & Sommerville, J. (2006). God does not play dice: Causal determinism and preschoolers' causal inferences. *Child development*, *77*(2), 427-442.

Seiver, E., Gopnik, A., & Goodman, N. D. (2013). Did she jump because she was the big sister or because the trampoline was safe? Causal inference and the development of social attribution. *Child development*, *84*(2), 443-454.

Sim, Z. L., & Xu, F. (2017). Learning higher-order generalizations through free play: Evidence from 2-and 3-year-old children. *Developmental psychology*, *53*(4), 642.

Sumner, E., Li, A. X., Perfors, A., Hayes, B., Navarro, D., & Sarnecka, B. W. (2019). The Exploration Advantage: Children's instinct to explore allows them to find information that adults miss.

Sumner, E., Steyvers, M., & Sarnecka, B. W. (2019). It's not the treasure, it's the hunt: Children are more explorative on an explore/exploit task than adults. *Proceedings of the cognitive sciences society 2019*, *41*.

Waismeyer, A., & Meltzoff, A. N. (2017). Learning to make things happen: Infants' observational learning of social and physical causal events. *Journal of experimental child psychology*, *162*, 58-71.

Walker, C. M., Bridgers, S., & Gopnik, A. (2016). The early emergence and puzzling decline of relational reasoning: Effects of knowledge and search on inferring abstract concepts. *Cognition*, *156*, 30-40.

Walker, C. M., & Gopnik, A. (2014). Toddlers infer higher-order relational principles in causal learning. *Psychological science*, *25*(1), 161-169.

Walker, C. M., & Gopnik, A. (2017). Discriminating relational and perceptual judgments: Evidence from human toddlers. *Cognition*, *166*, 23-27.

Xu, F., & Denison, S. (2009). Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition*, *112*(1), 97-104.

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences*, *105*(13), 5012-5015.