UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Constraints on Hypothesis Selection in Causal Learning

Permalink

https://escholarship.org/uc/item/3hj590td

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 37(0)

Authors

Tsividis, Pedro
Tenenbaum, Joshua B
Schulz, Laura

Publication Date

2015

Peer reviewed

Constraints on Hypothesis Selection in Causal Learning

Pedro Tsividis (tsividis@mit.edu), Joshua B. Tenenbaum (jbt@mit.edu), Laura Schulz (lschulz@mit.edu)
Massachusetts Institute of Technology, Department of Brain and Cognitive Sciences,
77 Massachusetts Ave., Cambridge, MA 02139 USA

Abstract

How do children identify promising hypotheses worth testing? Many studies have shown that preschoolers can use patterns of covariation together with prior knowledge to learn causal relationships. However, covariation data are not always available and myriad hypotheses may be commensurate with substantive knowledge about content domains. We propose that children can identify high-level abstract features common to effects and their candidate causes and use these to guide their search. We investigate children's sensitivity to two such high-level features — proportion and dynamics, and show that preschoolers can use these to link effects and candidate causes, even in the absence of other disambiguating information.

Keywords: Causal learning; intuitive theories; information search; analogy.

Introduction

The last fifteen years have produced a spate of research highlighting the kinds of epistemic practices that allow children to effectively navigate their complex world (see Gopnik & Wellman, 2012; Schulz, 2012b; Tenenbaum et al., 2011, for reviews.). Children rationally infer causal relationships from statistical evidence (e.g., Gopnik et al., 2004), selectively explore when evidence is confounded or surprising (Schulz & Bonawitz, 2007; Bonawitz et al., 2012), evaluate the relationship between samples and populations (Denison & Xu, 2010; Gweon et al., 2010; Xu & Denison, 2009), infer the existence of unobserved variables to explain anomalous data (Schulz et al., 2008), isolate candidate causes in order to distinguish between competing hypotheses (Cook et al., 2011; van Schijndel et al., 2015), and effectively search through hypothesis spaces for information (Nelson et al., 2014). These practices combine to enable the formation of *intuitive theories* — abstract, coherent, causal, ontologically-committed frameworks that guide prediction, explanation, and action (Gopnik & Meltzoff, 1997; Carey, 1985; Murphy & Medin, 1985).

Powerful as these kinds of epistemic practices are, they do not speak to how children might identify plausible hypotheses in the first place. Given that children do not always see a cause covary with an effect, and that there is an infinite space of hypotheses consistent with the learner's prior knowledge, children must have a method for constraining hypothesis spaces *before* engaging in (often costly) hypothesis-testing.

Langley et al. (1987) framed the process of scientific discovery as one of means-ends problem-solving, focusing on the value of heuristics that could enable scientists to reach a satisfying theory without exploring all possible alternative intermediate states. We suspect that child learners are guided by analogous heuristics for constraining their hypothesis spaces;

here we focus on whether children can use such heuristics to select between different hypotheses.

One recent proposal suggests that the problems considered by learners contain, in their abstract form, information about the abstract forms of their solutions (Schulz, 2012a; Magid et al., 2014). For instance, answers to "Where" questions are likely to involve two- or three-dimensional maps; answers to "When?" questions might involve timelines; answers to "Why" questions might involve chains or tree structures. In this sense, a well-posed problem already contains elements of its solution. "1917" might be the right or wrong answer with respect to when the Russian revolution took place, but with respect to the question of *why* it took place, it is not even in the space of possible answers.

Abstract information about the structural form of a solution may be especially important in the setting of causal learning. You may know, for instance, that you are trying to identify a causal mechanism responsible for something that blinks on and off. If you have a choice between a mechanism like a doorbell or a mechanism like a pulley, you might favor the former; given an effect whose outcome space is discrete, a candidate cause with discrete outcomes may seem preferable to one that has continuous outcomes. Of course, nothing guarantees that this inference is correct, but in the absence of other information, it is a reasonable strategy for narrowing down the hypothesis space. An initial test of this general idea showed that four- and five-year-olds were sensitive to abstract properties relating the form of candidate causes and effects. In a series of experiments, children successfully mapped discrete causes to discrete effects and continuous causes to continuous effects (Magid et al., 2014).

If priors about causal processes enable learners to select good hypotheses in the absence of covariation data or content-specific prior knowledge, what form should these priors take? If they are to be effective across a variety of problems, these priors cannot be about lower-level cognitive features such as color, height, pitch, and malleability, since causal relations do not often preserve these. That is, color changes are not usually caused by other colors; pitch changes are not caused by other pitch changes. More importantly, the space of possible mappings between specific low-level features is too large for such an approach to be efficient. However, certain higher-level amodal features such as extent, rate, arity (the number of states that a variable can take), distributional properties, and dynamics, are invariant to the lower-level features, allowing for more relevant and more efficient comparisons. We propose that children are sensitive to these higher-level features

and can use them to match effects with their causes. In this paper we focus on two: distributional properties and dynamics.

Experiment 1: Distributional Properties

As long as objects in two (or more) sets can be grouped into types, the relative proportions of those types across sets can be evaluated. This holds regardless of the features that serve to establish object identity, making such operations widely applicable. Even young infants have been shown to understand proportion, as they map proportion to probability of outcomes and can use proportion to guide their actions (Xu & Garcia, 2008; Xu & Denison, 2009; Denison & Xu, 2010). Here we ask whether young children believe that causal processes preserve proportion and if they can use this information to select between candidate causal hypotheses that cannot be distinguished by other means (e.g., covariation data, surface features, or domain-specific prior knowledge).

Participants Sixteen preschoolers (mean[1]: 5 years, 1 month; range: 4 years, 3 months — 5 years, 7 months.) were recruited from a local children's museum.

Materials We used Paint Tool SAI to create four flowers, two for each stimulus set. Within each stimulus set, the flowers differed in shape but matched in color (yellow for one stimulus set; blue for the other). For each stimulus set, there was a warm-up picture displaying only the two kinds of flowers and two test pictures: one test picture had 16 flowers of each kind (1:1 proportions); the other flower had 28 flowers of one kind and 4 of the other (7:1 proportions). We also used four different kinds of seeds, two for each stimulus set. Within each stimulus set, the seeds were near-identical to each other in size and texture, but different in color both from each other and the flowers (black and red seeds paired with yellow flowers in the first stimulus set, and brown and orange seeds paired with blue flowers in the second). The seeds were combined either in 1:1 or 7:1 proportions and were presented in containers, each containing approximately 100 seeds. See Figure 1.

Procedure Children were tested individually in a private room off the museum floor. The experimenter started the experiment by placing the warm-up picture of two flowers on the table in front of the child. He pointed to the two flowers and said, "Look, we have two flowers. This is a daisy and this is a lily. Now, you know how flowers are grown, right? With seeds! You put seeds into the ground and you water them and give them sun, and then flowers bloom! But seeds and flowers

---

[1]All children were 4 or 5 years old. Due to a data storage error, the ages of 8 of the children were only recorded accurately to the year; these children have been excluded from the estimated mean and range.
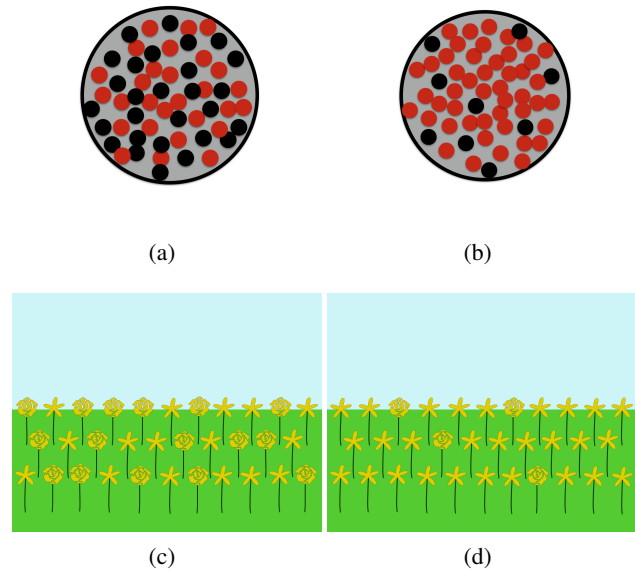

(a)   (b)

(c)   (d)

Figure 1: Schematic of seeds and actual flower fields used. Left: 1:1. right: 7:1.

are funny, because flowers end up not looking at all like the seeds they came from; the seeds change in all kinds of ways: they change in color, and size, and shape." Two seeds were placed on the table in front of the children, and introduced as the seeds that were used to make the flowers. Children were told, "These are the seeds we used to make the flowers. And just like we just talked about, they look totally different from the flowers, so we can't tell which seeds made which flowers just by looking at them." Then children were then told, "Now, we actually have whole fields of flowers, but before I show you the fields, let me tell you about how they were made." The two capfuls were brought out and placed on the table, next to each other (left-right randomized). "We had these two capfuls. And what we did was we took a bunch of seeds from this capful and threw them on one field, and we took a bunch of seeds from this capful and threw them onto the other field." The experimenter made a grabbing and throwing motion from each capful to the floor using alternating hands to illustrate. The experimenter then said, "Now I'll show you what the two fields ended up looking like," and brought out the two pictures of the fields of flowers (one with the 1:1 proportions and the other with the 7:1 proportions), placing them one above the other on the table (top-bottom randomized). He pointed to each field in turn and said, "Which capful do you think was used to make this field?".

After the child pointed to match each field with a capful, the experimenter removed all the stimuli and then repeated the procedure for a second trial with the second stimulus set, transitioning by saying, "Now, let me show you some more flowers." Presentation order of the fields and of the capfuls within stimulus set was randomized across and within participants, as was stimulus-set order.

Results There is no strict sense in which we can say that the children responded 'correctly' or 'incorrectly' given that there is no fact of the matter here. However, we can say whether children, as predicted, used the abstract property of proportionality to select one hypothesis over the other. Children were counted as having succeeded on a trial if they matched both capfuls in the trial to the correct fields. (The two questions — one for each field — within a trial were treated as non-independent because in introducing the capfuls, the experimenter had said that one capful was used for one field and one for the other. Thus the most conservative measure was to require a correct response to both questions). The probability of succeeding on both trials by chance is .25. Ten out of sixteen children responded at ceiling, answering correctly on both trials. The probability of success is .625; bootstrap-estimated 95% confidence interval: $[.354, .848]$. $p < .01$ by two-tailed binomial test. Figure 2 shows the number of children who were correct on 0, 1, or both trials.
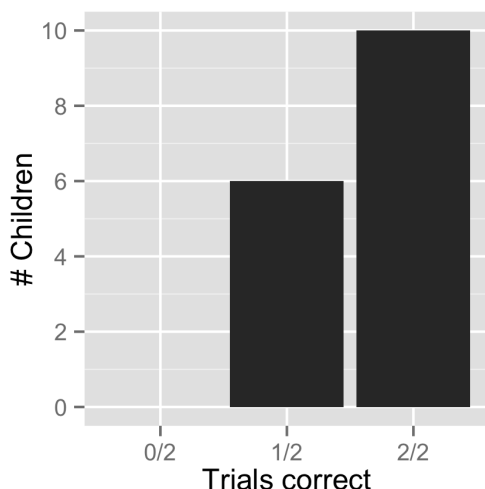

Figure 2: Experiment 1 results (N=16): Mean trials correct = 1.62. Probability of full success (2/2 trials) is .625; 95% confidence interval: $[.354, .848]$ (chance: .25). $p < .01$ by two-tailed binomial test.

Thus, children showed a clear preference for the proportion-preserving causal process, supporting our hypothesis.

Experiment 2: Dynamic Properties

Experiment 1 suggests that children are sensitive to proportionality in mapping causes to effects. However, if children have a general ability to identify plausible hypotheses using abstract amodal features, they should be sensitive to other kinds of relationships, as well. In Experiment 2, we look at children's sensitivity to dynamic properties. Specifically, we expect that, given data that saliently vary over time on some dimension, children will infer the latent structure of the variation and expect the cause of the data to possess similar latent structure. Children could be sensitive to a variety of dynamics. As a first pass, we investigate two: monotonicity and periodicity.

Participants Thirty-two preschoolers (mean: 4 years, 9 months; range: 4 years, 0 months — 5 years, 9 months.) were recruited from the local children's museum.

Materials We used sixteen 2"x2" pieces of white cardboard to make two sets of eight cards. For one set, four cards had a large red dot in the middle and four had a yellow one; for the other set, the dots on each of eight cards varied continuously from red through orange to yellow. These were presented in groups of eight, and represented the lights in two special rooms (See Figure 3). We also created four separate two-minute videos (two per stimulus set) using Adobe Flash and displayed them on the experimenter's laptop computer. Each video had ten identical 'alien bugs' moving around randomly. In two of these videos, the bugs changed in their speed over the course of the video. In the other two, the bugs grew spots on their back; the number of these spots changed throughout the course of the video. Each of these two features could change in two ways — either periodically or monotonically. We generated videos manifesting each of the four possible feature × dynamics combinations and split them into two stimulus sets as follows:

Stimulus Set 1 In video 1, the bugs' speed was governed by a periodic function — the bugs accelerated to a noticeably high speed (4 inches/second) over the course of 5 seconds, then decelerated to their original speed (.5 inches/second) over the course of 5 seconds, then re-accelerated to the high speed, and so on. This oscillation persisted throughout the two-minute video. In video 2, the bugs maintained a constant speed (.5 inches/second), but they grew black spots on their backs; the number of these spots increased from 0-20 over the course of the two-minute movie.

Stimulus Set 2 In video 1, the periodically-governed bugs moved at constant speed but varied in their number of spots, which oscillated between 0-10 spots (each half-period lasted 5 seconds, so that the fewest and most spots appeared matched the time-points at which the bugs were going slowest and fastest in stimulus set 1). In video 2, The monotonically-varying bugs changed in speed, starting out slowly and rising constantly throughout the video. See Figure 4 for a schematic depiction of these videos.

Procedure Children were told that they would be shown some alien bugs, but that before seeing the bugs, they would
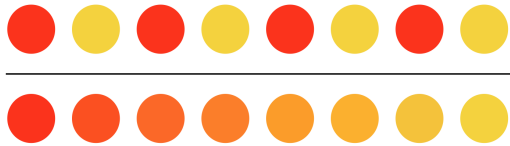
3

Figure 3: 'Lights' used for experiment 2. Top: periodic. Bottom: monotonic.

learn about the rooms the bugs were in. These rooms were described as identical except that they differed in their 'special lights'. The experimenter said, "The lights in the first room start out looking like this... then after a while they look like this... then after a while they look like this...", placing a light on the table each time he said 'this'. The lights were placed one by one on the table, left-to-right facing the child. Once the eight lights from the first set were placed, the experimenter said, "In the other room, the lights start out looking like this... then after a while they look like this..." and placed the lights for that room in the same manner as for the first. For one of the rooms the experimenter used the red and yellow (periodic) lights in alternation; for the other the experimenter used the continuously-varying red-to-yellow (monotonic) lights. Room type (periodic or monotonic) was randomized, as was whether the first light in each room was red or yellow. The language used to describe both sets of lights was identical. The first four lights in each set were placed on the table approximately every 3 seconds; to keep the description conversational, each of the last four lights was placed approximately every 1 second. This also ensured that there was no way to map the rate of presentation of the cards to the rate of change of either speed or spots in either display.

Following this, the children were invited to look at the bugs. In the first stimulus set, for the periodic bugs, the children were invited to attend to their speed: as the video played, the researcher pointed out when the bugs sped up ("See, now they're getting faster") and when they slowed down ("... and now they're getting slower"). For the monotonic bugs, the video was played twice. The first time, the experimenter allowed the children to observe the changing number of spots on their own. After there were approximately 15 spots on each bug, the experimenter restarted the movie, and this time counted the number of spots on the bugs as these increased in number, summarizing the change after there were 6 spots ("They're getting more and more spots.").

For the second stimulus set, the procedure was identical, except that the language was modified appropriately to describe the different changes in the bugs. The experimenter pointed out when the spots on the periodic bugs were increasing or decreasing in number ("Now they're getting more/fewer spots"), and on the monotonic bugs, pointed out the increasing speed ("Now they're going faster... and now the're going faster... and now they're going even faster...",
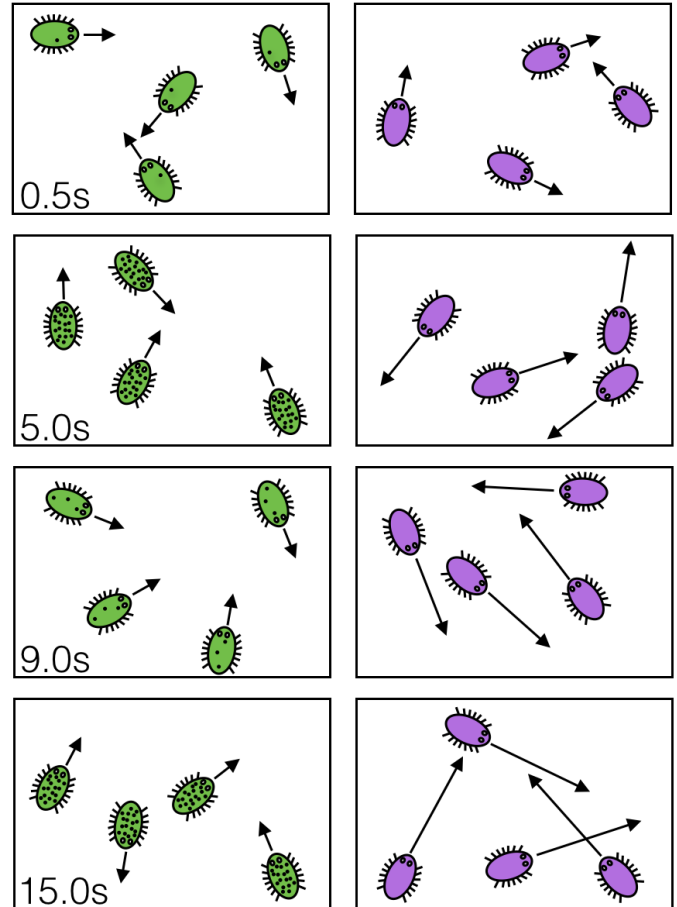

Figure 4: Periodic (green) and monotonic (purple) bugs from stimulus set 2. The green bugs move at constant speed but increase and decrease in number of spots; the purple bugs increase speed monotonically, as indicated by the increasing length of the vectors.

etc.).

Each child only saw one stimulus set — that is, one set of periodic bugs and one set of monotonic bugs. Stimulus-set assignment was counterbalanced.

After the children were familiarized with the stimuli, they were shown the first bugs they had seen, asked to remember how they changed, and then were told the following: "So, we know that these guys are in one of these two rooms that we talked about before. They can see the lights in the room but we can't. And what's causing the speed to change is the lights of the room they're in. They could be in this room or in this room. Do you know what room they're in?" Children were asked to point to the room they thought the bugs were in. After this, they were shown the other bugs and the above description and question were repeated verbatim, changing only 'speed' to 'spots' (or vice-versa if the first-seen bugs had varied in spots). It was emphasized that each of the bugs could be in each of the two rooms.

4

2437

Results Due to the fact that each group of bugs was independently described as being potentially in each room, the two questions are independent; the probability of answering both questions correctly by chance is .25. Children were sensitive to the fact that bugs could have been in the same room, as evidenced by the fact that six children placed both types of bugs in the same room. Nineteen out of thirty-two children were at ceiling, answering correctly on both trials. The probability of success is: .594; bootstrap-estimated 95% confidence interval: $[.406, .763]$. $p < .001$ by two-tailed binomial test. Figure 5 shows the number of children who were correct on 0, 1, or 2 trials. Performance across the two stimulus sets was comparable: 10/16 children were at ceiling for the first and 9/16 children were at ceiling for the second.
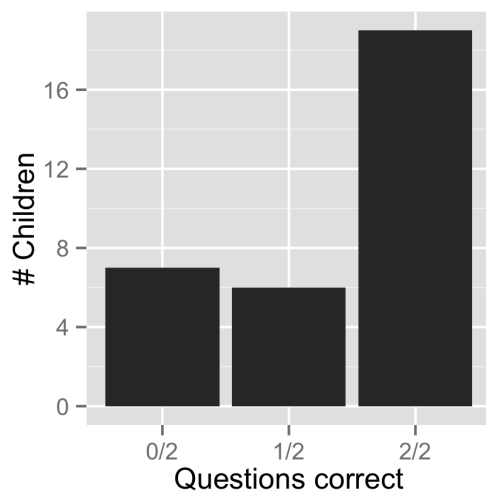

Figure 5: Experiment 1 results (N=32): Mean trials correct = 1.375. Probability of full success (2/2 trials): .594, 95% confidence interval: $[.406, .763]$ (chance: .25). $p < .001$ by two-tailed binomial test.

Discussion

In two experiments, we investigated whether children use higher-level features of data to match causes to effects. In Experiment 1, these features were static and distributional: children showed a preference for a causal process that preserved distributional identity, matching 1:1 seeds to 1:1 flowers and 7:1 seeds to 7:1 flowers irrespective of surface properties of the seeds and flowers (size, color, texture, etc.). In Experiment 2, the higher-level features were dynamic: children showed a preference for a causal process that preserved dynamic form — periodic or monotonic — irrespective of the lower-level features in which these dynamics were manifested; periodically-varying lights were seen as the cause of the spots on bugs *or* the speed of the bugs, depending on which one varied periodically, and monotonically-varying

lights were seen as the cause of the monotonically-varying feature of bugs. In both experiments, children received no information about how causes and effects covaried; inferences were made based only on abstract properties of the stimuli.

The idea of using high-level features to match percepts has been presented previously, in the literature on cross-modal matching (see, e.g., Lewkowicz & Turkewitz, 1980; Spence, 2011). Cross-modal matching is often presented as a partial solution to the binding problem and the focus is on mappings between stimuli belonging to different perceptual modalities (i.e., the sight and feel of a stimulus). We believe that it is possible that the inferences children drew here and those shown in studies on cross-modal matching may rely on the same representational machinery; specifically, we believe that both rely on the use of particularly powerful higher-level features of the stimuli. However, the mechanisms we suggest are applicable to a far wider array of problems than mere cross-modal matching. The higher-level features we have examined, namely distributional properties and dynamics, are calculable both within and across modes; in our case we have examined their application within a perceptual modality and have found them to apply both to naturalistic stimuli such as seeds and flowers and to arbitrary stimuli, such as randomly-moving animated alien bugs. In principle, the same kinds of inferences could be used for problems entirely abstract in nature (e.g., using the dynamics of the interest rate to map it onto changes in the Gross Domestic Product).

Research on analogy (Gentner & Markman, 1997; Gentner, 1977; Gick & Holyoak, 1980) has shown that children and adults are able to bring distinct mental representations into structural alignment and to use the relations that obtain within one domain to reason about the other. We believe that analogical reasoning is an elegant example of the more general ability to use high-level features to constrain hypothesis spaces. Note that here, however, we did not set up a situation where children could go from a known problem and solution to a new problem and a new solution by setting up a relational mapping between arguments (e.g., Christie & Gentner, 2010). Rather, children had to infer the representation that might connect the form of the effect to the form of the candidate causes and use this representation to guide their responses. We emphasize this not to minimize the importance of analogical reasoning, but because the general ability to represent these abstract high-level predicates may allow learners to narrow the hypothesis space even when problems do not present as analogies.

A variety of empirical questions remain: We have shown that children are sensitive to high-level features like proportionality and periodic or monotonic dynamics, and can use these to infer causal relationships. Do children in fact use these features to constrain hypothesis generation, or to constrain a large hypothesis space? When such features conflict with lower-level features, such as color, texture, or size, how do children resolve the conflict? Perhaps most interestingly,

how do children learn these high-level predicates, and which predicates are available to them at different times throughout development? Given the prevalence of phenomena in the world for which distributional properties and dynamics are coherent and relevant, it may not be surprising that four- and five-year-olds can use these; what about younger children and what about other kinds of properties (e.g.,. extent, other kinds of dynamics, richer distributional information, or combinations of these)? Much remains to be understood about how children identify abstract features, and about what other kinds of high-level features they can recognize; much also remains to be understood about how we might computationally characterize the ability to represent, learn, and use these higher-level features. We hope to address these questions in future work.

Acknowledgments

References

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*(4), 215–234.

Carey, S. (1985). *Conceptual change in childhood*. MIT press.

Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development*, *11*(3), 356–373.

Cook, C., Goodman, N. D., & Schulz, L. (2011). Where science starts: Spontaneous experiments in preschoolers exploratory play. *Cognition*, *120*(3), 341–349.

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.

Gentner, D. (1977). Children's performance on a spatial analogies task. *Child development*, 1034–1039.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American psychologist*, *52*(1), 45.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive psychology*, *12*(3), 306–355.

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

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. MIT Press.

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

Gweon, H., Tenenbaum, J. B., & Schulz, L. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences*, *107*(20), 9066–9071.

Langley, P., Simon, H., Bradshaw, G., & Zyikow, J. (1987). *Scientific discovery: Computational explorations of the creative processes*. MIT press.

Lewkowicz, D. J., & Turkewitz, G. (1980). Cross-modal equivalence in early infancy: Auditory–visual intensity matching. *Developmental psychology*, *16*(6), 597.

Magid, R., Shulz, L., & Sheshkin, M. (2014). Imagination and the generation of new ideas.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological review*, *92*(3), 289.

Nelson, J. D., Divjak, B., Gudmunsdottir, G., Martignon, L. F., & Meder, B. (2014). Children's sequential information search is sensitive to environmental probabilities. *Cognition*, *130*, 74–80.

Schulz, L. (2012a). Finding new facts; thinking new thoughts. In F. Xu, T. Kushnir, & J. Benson (Eds.), *Advances in child development and behavior*. Oxford: Academic Press.

Schulz, L. (2012b). The origins of inquiry: Inductive inference and exploration in early childhood. *Trends in cognitive sciences*, *16*(7), 382–389.

Schulz, L., & Bonawitz, E. B. (2007). Serious fun: preschoolers engage in more exploratory play when evidence is confounded. *Developmental psychology*, *43*(4), 1045.

Schulz, L., Goodman, N. D., Tenenbaum, J. B., & Jenkins, A. C. (2008). Going beyond the evidence: Abstract laws and preschoolers responses to anomalous data. *Cognition*, *109*(2), 211–223.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971–995.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *science*, *331*(6022), 1279–1285.

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.

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.

6