Psychological Science, In Press

Running Head: 18-30-MONTH-OLDS INFER HIGHER-ORDER RELATIONS

18-30-Month-Olds Infer Higher-Order Relational Principles in Causal Learning

Caren M. Walker & Alison Gopnik

University of California, Berkeley

Department of Psychology

3210 Tolman Hall, Berkley, CA 94720

Author Note

Correspondence concerning this article should be addressed to Caren Walker, Department of Psychology, University of California, Berkeley, 3210 Tolman Hall, Berkeley, California 94720. Email caren.walker@berkeley.edu

Figures = 4; word count = 3,997

Abstract

Children make inductive inferences about the causal properties of individual objects from a very young age. When can they infer higher-order relational properties? In three experiments, we examined 18- to 30-month-olds' relational inferences in a causal task. Results suggest that these children are able to infer a higher-order relational causal principle from just a few observations and use this inference to guide their own subsequent actions and bring about a novel causal outcome. Moreover, they passed a version of the relational task that is very difficult for non-human primates. Findings are considered in light of recent discussion about the nature of relational and causal reasoning, and their evolutionary origins.

KEYWORDS: cognitive development, causality, infant development, inference, learning

18-30-Month-Olds Infer Higher-Order Relations in Causal Learning

Learning about causal relationships is one of the most important and challenging problems young humans face. Causal knowledge allows you to act on the world – if you know A causes B, you can act on A to bring about B. Studies show that children as young as 16 to 24 months of age can quickly learn causal properties of objects from patterns of statistical contingency and can act on that knowledge to bring about effects (e.g., Sobel & Kirkham, 2006; Meltzoff, Waismeyer & Gopnik, 2012; Gweon & Schulz, 2011; for reviews see Gopnik, 2012; Gopnik & Wellman, 2012).

Much of this research on early causal learning has used a "blicket detector" paradigm (Gopnik & Sobel, 2000), in which children learn which objects activate a novel machine. Children's inferences in these tasks go beyond associative learning, revealing the distinctive profile of causal inference. For example, children will use these inferences to design novel interventions – patterns of action they have never observed – to construct counterfactuals and make explicit causal judgments, including judgments about unobserved features (e.g. Gopnik & Sobel, 2000; Gopnik et al. 2004; Schulz, Gopnik, & Glymour, 2007; Sobel, Yoachim, Gopnik, Meltzoff, & Blumenthal, 2007).

However, we know much less about the development of children's ability to infer higher-order relational causal principles. According to "theory theorists" of cognitive development, children are not only learning particular causal relationships, but also higher-order generalizations *about* causal structure (e.g., Carey, 2010, Wellman & Gelman, 1992, Gopnik & Meltzoff, 1997). Recent computational work also suggests that higher-order generalizations can help children learn new specific relationships from perceptual data more quickly (e.g., Griffiths

& Tenenbaum, 2007; Goodman, Ullman, & Tenenbaum, 2011; Kemp, Perfors & Tenenbaum, 2007).

Causal inferences might be more or less abstract, higher-order, or relational in many different ways. Here we focus on just one contrast: between object properties, such as shape or color, and higher-order relations between those properties, such as whether they are the same or different. For example, very young children can learn that red blocks activate a toy. When can they learn that two blocks that are the same (regardless of their color) can do so?

Empirical research using looking-time measures suggests that human infants may be able to recognize patterns of data that involve higher-order relations such as "same" (Dewar & Xu, 2010; Ferry, Hespos & Gentner, 2012; Tyrell, Stauffer & Snowman, 1991). However, there is no evidence to date that infants can use those patterns to make causal inferences or guide subsequent actions.

In fact, earlier studies indicate that even preschoolers have difficulty making inferences in higher-order relational reasoning tasks (e.g., Christie & Gentner, 2007, 2010; Gentner, 2010). Children succeeded only when given labels or linguistic scaffolding to point out the pattern of similarity. Indeed, even when explicitly instructed to compare objects, 3-year-olds' performance was tenuous, dropping significantly below chance when test items were presented sequentially, rather than simultaneously (Christie & Gentner, 2010).

These findings might lead to the conclusion that learning higher-order relations and using them to guide actions depends on direct instruction, language, and cultural input (e.g., Christie & Gentner, 2007; Gentner, 2003, 2010; Gentner, Anggoro, & Klibanoff, 2011). However, these tasks often relied on verbal categorizations of complex, multi-dimensional stimuli (e.g., Christie & Gentner, 2010). One study by Smith (1984) provides a hint that children might do better in a more goal-directed task with simpler materials. In particular, 2½-year-olds showed some understanding of identity matching in a non-verbal game.

Higher-order relational reasoning has also been studied extensively in non-human animals. Chimpanzees, like young infants, are able to spontaneously detect a relational pattern in habituation tasks (Oden, Thompson, & Premack, 1990). However, they have more difficulty with a relational match-to-sample task (Premack, 1983; Oden, Premack, & Thompson, 1988). In these tasks, animals observe a relational pattern – AA', BB', and CC' all lead to a reward. Then they are given a choice between AB (object match) and DD' (relational match). Although A and B have each been associated with the reward, an animal who has inferred the higher-order relational pattern should choose DD'. Premack and colleagues have found that chimpanzees could not solve this relational task without hundreds of trials with feedback (Premack, 1988) or training to use linguistic symbols for "same" (Premack, 1976; 1983; Premack & Premack, 1983; 2002).

Additional comparative studies confirm that this task is especially difficult for non-human primates and other animals (see Penn, Holyoak, & Povinelli, 2008). Moreover, when non-human animals, such as baboons, *do* solve this task, they require extended training and thousands of trials, which may indicate the use of simpler perceptual strategies such as minimizing entropy in a perceptual array (Fagot, Wasserman, & Young, 2001; Wasserman, Fagot, & Young, 2001).

Do human children always require linguistic cues or extensive training to solve relational tasks like the preschoolers and primates in earlier studies? We designed a non-verbal "blicket detector" task to explore when children could use higher-order relations to make causal inferences. In contrast to previous studies of causal inference, the causal effect depended on whether the objects were the same or different, rather than on properties of the objects themselves.

Experiment 1

In Experiment 1, 21- to 24-month-olds were introduced to a novel toy that played music and 3 unique pairs of identical blocks: AA', BB', and CC'. The experimenter placed blocks on the toy and the toy either activated or did not. Although individual blocks failed to activate the toy alone, pairs of identical blocks produced the effect. Immediately after this brief training, we examined whether children learned the novel relational property (i.e., "same") by asking them to activate the toy.

Methods

Participants

A total of 23 21- to 24-month-old toddlers participated in Experiment 1 (*M* = 23.0 months; *SD* = 1.05 months; range = 20.9-25.0 months; 13 girls). Three additional children were tested but excluded for fussiness or for failing to respond. Children were recruited from daycare centers and museums, and a range of ethnicities resembling the diversity of the population was represented.

Materials

The toy was a 10" x 6" x 4" opaque white cardboard box containing a wireless doorbell. When a block "activated" the toy, the doorbell played a melody. In fact, the toy was surrepticiously activated by a remote control. Six painted wooden blocks in assorted colors and shapes (3 unique pairs of 2 identical blocks) were placed on the toy during the training phase. Six additional blocks were used during the test phase, including 2 novel pairs of identical blocks and 2 novel individual blocks.

Procedure

The procedure for Experiment 1 is illustrated in Figure 1. Following a warm-up, the toy was placed on the table. The experimenter said, “This is my toy. Some things make my toy play music and some things do not make my toy play music.” Children then observed while the experimenter placed 6 blocks (3 unique pairs of “same” objects: AA’, BB’, CC’) on the table in front of the toy. She said, “Let’s try one!”, selected a block (A) and placed it on top of the toy. No effect was produced. After a pause, the experimenter again said, “Let’s try one!”, selected the paired block (A’) and placed it next to the first block (A) on top of the toy. This pair of objects (AA’) activated the toy. The experimenter smiled and said, “Music!”, removed the blocks and returned them to the pile of 6. This procedure was repeated with the two remaining pairs (BB’ and CC’). The order of the pairs was randomized. Following all three demonstrations, all blocks were removed.

Next, the experimenter produced 3 test blocks (1 *novel paired block* (D), 1 *familiar block* (A), and 1 *novel distractor block* (E) and placed them in a row on the table. The order of presentation was randomized. She said, “Let’s try one!”, produced the *target block* (D’), and placed it on top of the toy. No effect was produced. The experimenter then pushed the toy and all 3 test blocks towards the child, and asked, “Can you pick one of these (pointing to the test blocks) to make my toy play music?”

The first test block that the child placed on the toy was recorded. The toy activated if the child correctly selected the *novel paired block* (D). If the child selected the *familiar block* (A) or the *novel distractor block* (E), the toy failed to activate. After this feedback, this procedure was repeated in a second test trial with a new set of test blocks.

If toddlers were acting based on the previous association between the block and effect, they should choose the *familiar block* (A). If they simply preferred to try novel blocks they should pick the *novel distractor block* (E) as often as the *novel paired block* (D). However, if toddlers were able to learn the higher-order relation, they should select the *novel paired block* (D).

Coding & Reliability

Children received 1 point for selecting the *novel paired block* and 0 points for selecting either of the other blocks in each trial. Responses were recorded by a second researcher during the testing session, and all sessions were recorded for independent coding by a third researcher who was naïve to the the hypotheses of the experiment. Interrater reliability was very high; the two coders agreed on 99% of the children's responses. Two minor discrepancies were resolved by a third party.

Results & Discussion

Across the two test trials, children inferred the relational property and selected the *novel paired block* (D) more often than expected by chance ($M = 1.13$, $SD = .82$), $\chi^2(2) = 19.07$, $p < .001$. Fischer exact test revealed no order effects for test trials 1 and 2, $p = .39$. Children chose the *novel paired block* (61%) significantly more often than the *novel distractor block* (20%), $\chi^2(2) = 14.15$, $p < .001$ and significantly more often than the *familiar block* (15%), $\chi^2(2) = 14.09$, $p < .001$. A minority of children (4%) placed more than one block on the toy simultaneously, and were scored as incorrect.

Previous proposals have suggested that children are unable to reason relationally because they tend to focus on the identity of objects that have been previously associated with the outcome (e.g., Gentner, 2010). We show no evidence of this. In fact, only 39% of participants who answered incorrectly on a given trial selected the *familiar block*, with no difference in their selection between the *familiar block* and the *novel distractor*, $\chi^2(2) = 2.43$, $p = .30$. This is particularly surprising, given that this block had been associated with the effect during training.

Results suggest that by 21-24 months of age, toddlers are able to infer the causal principle – "same" – from just a few pieces of evidence and use this inference to bring about a novel causal outcome. However, children might have succeeded on this task by imitating the experimenter's selection or because they preferred to match, regardless of training. Experiment 1a was designed to address these alternatives.

Experiment 1a

The procedure for Experiment 1a was identical to Experiment 1, but the second object in the pair was occluded. Because children only observed the first item in each pair, they were given no evidence for the relational property. If children were simply imitating the experimenter or had a preexisting preference for matching, then children's performance should not differ from Experiment 1.

Methods

Participants

Twenty 21-24-month-olds participated (*M* = 22.4 months; *SD* = 1.8 months; range = 20.8-25.6 months; 8 girls). Two additional children were tested but excluded for failing to respond. Recruitment procedures and demographics were identical to Experiment 1.

Materials & Procedures

Materials and procedures were identical to Experiment 1. However, children did not observe the second object during the training trials. Instead, the second object was occluded by a 4" x 4" square piece of cardboard. Additionally, only one test trial was administered in order to avoid providing feedback. Therefore, children could receive 0 or 1 points. Interrater reliability for Experiment 1a was 100%.

Results & Discussion

In the absense of evidence for the relational principle, only 40% of participants selected the paired block, [exact binomial test, $p = .65$, ns], which was significantly different from the percentage of children (61%) of the same age on their first trial in Experiment 1, $p < .05$ by Fischer's exact test (see Figure 2). Childrens' selections were evenly distributed: 40% of children selected the *novel paired block,* 35% of children selected the *familiar block*, and 25% of children selected the *novel distractor*. These results show that the findings from Experiment 1 could not have been the result of imitation or a bias to match.

Experiment 2

In the earlier primate studies, the canonical *relational* match to sample tasks presented pairs simultaneously during training (e.g., the relation "same" was taught using pairs AA' and BB'), and the animals had to choose between test pairs illustrating "same" (CC') and "different" (DE). Chimpanzees were unable to spontaneously succeed on this task – and had great difficulty even after engaging in trial and error over hundreds of trials. However, chimpanzees *were* able to solve a simpler match-to-sample task. In these tasks, the animals were first taught to match a test object (A) to a target object (A') through multiple positive and negative reinforcement trials over several weeks. They then generalized this pattern to novel objects without additional training (Premack, 1976; Premack & Premack, 1983; 2003; Oden et al, 1988).

Our task in Experiment 1, like the simple primate match to sample task, presented the training objects sequentially, and this may have made the task easier. However, Experiment 1 also differed in several ways from the primate task. Children learned by observation – they did not initially make the responses themselves – and they spontaneously chose the *novel paired block* after observing only three trials. Additionally, they never observed that the mismatching block would "not" produce the effect, so the association between the incorrect *familiar* block and the effect should have continued to be high.

In order to make the comparison to the primate tasks clearer, we designed a causal task that was more directly analogous to the primate relational match to sample tasks, in which both "same" and "different" objects are presented in pairs. This task also allowed us to explore whether children would infer the "different" relation as well as "same." We included toddlers from a broader age range to explore possible developmental differences, recruiting children aged 18 to 30 months.

Participants were randomly assigned to one of two conditions: *same* or *different.* In the *same* condition, children were given two pieces of evidence that pairs of "same" objects (AA', BB') simultaneously placed on the toy produced the effect. We also provided two pieces of evidence that pairs of "different" objects (DE, FG) *failed* to produce the effect. In the *different* condition, children were given the same four pieces of evidence but "different" pairs (DE, FG) produced the effect, while "same" pairs (AA', BB') *failed* to do so.

Methods

Participants

Thirty-eight 18-30-month-olds participated (M = 25.8 months; SD = 3.8 months; range = 18.0-30.6 months; 21 girls), with 19 children randomly assigned to each condition (*same* and *different*). Seven additional children were tested but excluded: 4 due to failure to complete the study and 3 due to experimenter error. Recruitment procedures and demographics were identical to Experiments 1-1a.

Materials

The same toy from Experiments 1 and 1a was used. Eight painted wooden blocks in assorted colors and shapes (2 pairs of “same” blocks and 2 pairs of “different” blocks) were placed on the toy in pairs during training. The “same” blocks were identical in color and shape, and the “different” blocks were distinct in color and shape (see Figure 3). Four additional blocks were used during the test phase, including 1 novel pair of “same” blocks and 1 novel pair of “different” blocks. The pairs of test blocks were placed on 4” x 4” plastic trays.

Procedure

The procedure for Experiment 2 is illustrated in Figure 3. Following a warm-up, the toy was placed on the table. The experimenter said, “This is my toy. Some things make my toy play music and some things do not make my toy play music.” Children then observed while the experimenter placed all 8 training blocks (A, A’, B, B’, E, F, G, H) in a random arrangement on the table in front of the toy. The experimenter said, “Look at these things! We will try them on my toy.” Then, the experimenter removed all objects from view, selected the first pair of blocks (e.g., AA’), and placed the blocks simultaneously on the toy. Children in the *same* condition observed the pair of “same” objects activate the toy. The experimenter smiled and said, “Music! Let’s try that again!”, picked up the pair of blocks, and placed them back on the toy a second time, and children again observed the outcome. After this second demonstration, the experimenter removed the pair, selected another pair – a “different” pair (e.g., EF) – and placed it on the toy. This time, children in the *same* condition observed no effect. As with the first pair, this was demonstrated a second time before moving on to the third pair. This procedure was repeated for all 4 pairs: 2 pairs of “same” objects and 2 pairs of “different” objects. All pairs were placed on the toy twice. Therefore, children observed a total of 8 outcomes (4 positive and 4 negative).

Children in the *different* condition observed the same set of evidence as children in the *same* condition, with one critical change: pairs of “different” objects (e.g., EF) caused the toy to play music, while the pairs of “same” objects (e.g., AA’) failed to activate the toy. There were no other differences in procedure. The particular objects included in each pair was randomized, as well as the order that the pairs were presented during training.

Following the training phrase in both conditions, the experimenter said, “Now it is going to be *your* turn. I want you to help me pick the ones that will make my toy play music!” The experimenter produced 2 pairs of test blocks (1 novel “same” pair [JJ’], 1 novel “different” pair [KL]). In order to avoid a novelty preference, *both* test pairs were composed of novel objects. The pairs were presented to the child on plastic trays. The experimenter held up the two trays, shook them to get the child’s attention, and asked, “Can *you* pick the ones that will make my toy play music?” She then placed the trays on opposite sides of the table in front of the child. The side on which the correct pair was placed was randomized between subjects. The first tray that the child selected was recorded. Correct selections included pointing to the tray, reaching to the tray, or picking up the objects on the tray.

If they learned the relational property, then children in the *same* condition should correctly select the tray with the novel “same” objects (AA’), while children in the *different* condition should correctly select the tray with the novel “different” objects. Correct selections were given a score of “1” and incorrect selections were given a score of “0”. Coding and recording procedures were identical to Experiments 1-1a. Interrater reliability was very high; the two coders agreed on all but one of the children’s responses to the test questions.

Results & Discussion

Results of Experiment 2 appear in Figure 4. Children inferred the relational property and selected the correct pair more often than expected by chance (M = .79, SD = .41; chance = .5), [exact binomial test], p<.02] in both *same* and *different* conditions. In fact, performance was identical in the *same* and *different* conditions, with 15 out of 19 children in each condition selecting the test pair that corresponded with the relation learned during the training trials. Additionally, logistic regression revealed no significant developmental change in performance between 18 and 30 months of age, $\chi^2(1, 38) = .11$, $p = .74$ (ns). The fact that children responded differentially in the otherwise identical *same* and *different* conditions also allowed us to rule out superficial explanations for the results, such as imitation or a preference for same or different pairs – each condition acted as a control for the other condition.

Experiment 2 indicates that toddlers are able to infer the relational causal principles "same" and "different" from a just few pieces of evidence, and use this inference to intervene to bring about a novel causal outcome.

General Discussion

These findings show that human toddlers as young as 18 months can succeed on a causal relational match-to-sample task after only a few trials and without explicit linguistic cues, instruction, or reward. This study has implications for our understanding of both causal and relational reasoning. Using this method, toddlers are able to quicky learn higher-order relational causal principles and use them to guide their actions. This ability appears to be in place surprisingly early -- only a few months after the first evidence of the ability to learn about specific causal properties from contingency -- and it may be in place even earlier. This may help explain how children acquire the impressive causal knowledge evident in early "intutive theories" (Gopnik & Wellman, 2012; Carey, 2010).

These findings also contrast with the striking failure of non-human primates to solve similar tasks, even when the relation is associated with a strong pattern of positive and negative reinforcement, and even after hundreds (or thousands) of trials. This finding might support the suggestion that an ability to quickly learn relational causal concepts is a dimension on which humans differ from other primates. This might in turn reflect the broader evolution of higher-order relational cognition (Penn, Holyoak & Povinelli, 2008) or causal cognition in general (Heyes & Frith, 2012; Byrne, 1995; Buchsbaum et al., 2012).

Several questions for further study remain. One is whether the causal nature of this task was critical, or whether other aspects of the task, such as the fact that it involved goal-directed actions, might have made it easier for the children than relational tasks in other studies. It is also possible that children could succeed on this particular task by basing their causal inference on the observed association between the higher-order relational features and the effects. In other "blicket detector" studies children's inferences go beyond association, but those studies would have to be replicated with the current relational design.

Further, it is possible that the children's success was due to a perceptual heuristic, as has been suggested for non-human primates (Penn et al., 2008; Fagot, et al., 2001; Wasserman, et al., 2001). According to this argument, it is possible to solve relational match-to-sample tasts using the perceptual cue, entropy (i.e., the Shannon entropy of AA' is 0, while that of AB is 1). Several features of the children's behavior weigh against this possibility: children saw pairs of objects (rather than multi-element displays), they observed only two positive and two negative trials, they never acted on the object, and their behavior was never reinforced. Indeed, no other species has come close to demonstrating the first-trial performance of these human children after so few observations (see Penn, et al., 2008). Additionally, although human participants have been shown to be sensitive to entropy, findings suggested that additional processes of categorization likely play a role in the human conceptualization of “same-different” relations (Fagot, et al., 2001). Nevertheless, future research examining this possibility would be informative.

Finally, it will be important to replicate this particular task with non-human primates to determine if, like children, they show greater success, or continue to have difficulty. Our protocol did not require a verbal response, so it may be useful in examining reasoning capacities in both pre-verbal human infants and possibly in non-human animals.

However, the current study does suggest that the ability to infer causal higher-order relations, an ability which could play a crucial role in further learning, is in place in humans from a very early age and does not depend on explicit linguistic cues or cultural scaffolding.

Acknowledgments

Research was funded by the James S. McDonnell Foundation and the National Science Foundation (BCS-1023875) to A. Gopnik. We thank the parents and children who participated and the University of California, Berkeley Early Childhood Centers, Lawrence Hall of Science, and Habitot. We are grateful to Fei Xu, Tania Lombrozo, Keith Holyoak, and anonymous reviewers for their feedback, to Rosie Aboody, Anna Akullien, Ela Banerjee, Sierra Eisen, Brynna Ledford, and especially Samantha Hubachek for facilitating data collection.

References

Buchsbaum, D., Bridgers, S., Weisberg, D.S., & Gopnik, A. (2012). The power of possibility: Causal learning, counterfactual reasoning, and pretend play. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367,* 2202-2212.

Byrne, R.W. (1995). *The thinking ape: Evolutionary origins of intelligence.* New York: Oxford University Press.

Carey, S. (2010). *The origin of concepts*. USA: Oxford University Press

Christie, S. & Gentner, D. (2007). Relational similarity in identity relation: The role of language. In Vosniadou, S. & Kayser, D. (Eds.). *Proceedings of the Second European Cognitive Science Conference.* London, UK: Taylor & Francis.

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

Dewar, K.M., Xu, F. Induction, overhypothesis, and the origin of abstract knowledge: Evidence from 9-month-old infants. *Psychological Science. 43*, 1227-1238.

Fagot, J., Wasserman, E.A., & Young, M.E. (2001). Discriminating the relation between relations: The role of entropy in abstract conceptualization by baboons (*Papio papio*) and humans (*Homo sapiens*). *Journal of Experimental Psychology, 27*(4): 316-328.

Ferry, S. Hespos, S. Gentner, D. (2012). Prelinguistic relational concepts: Investigating the origins of analogical reasoning in infants. Poster presented at the *18th Biennial International Conference on Infant Studies*, in Minneapolis, MN, June 7, 2012.

Gentner, D. (2003). Why we're so smart. *Language in mind: Advances n the study of language and thought,* 195-235.

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems.

Cognitive Science, 34, 752-775.

Gentner, D. Anggoro, F.K., Klibanoff, R.S. (2011). Structure mapping and relational language support children's learning of relational categories. *Child Development, 82*, 1173-1188.

Goodman, N. D. (1955). *Fact, fiction, and forecast.* Cambridge, MA: Harvard University Press.

Gopnik, A. (2012). Scientific thinking in young children: Theoretical advances, empirical research, and policy implications. *Science, 337*, 1623-1627.

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

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,* 3–32.

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

Gopnik, A. & Wellman, H. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin, 138*, 1085-1108.

Griffiths, T., & Tenenbaum, J. B. (2007). Two proposals for causal grammars. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 323–345). New York, NY: Oxford University Press.

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

Heyes, C. & Frith, U. (Eds.) (2012). New thinking: The evolution of human cognition (Theme Issue). *Philosophical Transactions of the Royal Society, B,* 367.

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

Meltzoff, A. Waismeyer, A. Gopnik, A. (2012). Learning about causes from people: Observational causal learning in 24-month-olds infants. *Developmental Psychology,* Online First Publication.

Oden, D. L., Premack, D., & Thompson, R. K. (1988). Spontaneous Transfer of Matching by Infant Chimpanzees (Pan-Troglodytes). *Journal of Experimental Psychology-Animal Behavior Processes, 14(2)*, 140-145

Oden, D.L., Thompson, R.K., Premack, D. (1990). Infant chimpanzees spontaneously perceive both concrete and abstract same/different relations. *Child Development, 61*(3): 621-31.

Penn, D.C., Holyoak, K.J., Povinelli, D.J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences, 31,* 109-178.

Premack, D. (1976). *Intelligence in ape and man*. Hillsdale, N.J.: L. Erlbaum Associates

Premack, D. (1983). The codes of man and beasts. *Behavioral & Brain Sciences, 6,* 125-167.

Premack, D. (1988). Minds without language. In L. Weiskrantz, (Ed.), *Thought without language* (pp. 46-65). Oxford: Clarendon Press.

Premack, D. & Premack, A.J. (1983). *The mind of an ape*. New York: W.W. Norton.

Premack, D. & Premack, A.J. (2002). *Original intelligence: Unlocking the mystery of who we are.* New York: McGraw-Hill.

Premack, D. & Premack, A. J. (2003). *Original intelligence: unlocking the mystery of who we are.* New York: McGraw-Hill.

Schulz, L., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science (special section on Bayesian and Bayes-Net approaches to development)* 10(3):322-332.

Smith, L.B. (1984). Young children's understanding of attributes and dimensions: A comparison of conceptual and linguistic measures. *Child Development, 55*(2): 363-380.

Sobel, D. Kirkham, N.Z. (2006). Blickets and babies: The development of causal reasoning in toddlers and infants. *Developmental Psychology, 42*, 1103-1115.

Sobel, D. M., Yoachim, C. M., Gopnik, A., Meltzoff, A. N., & Blumenthal, E. J. (2007). The blicket within: Preschoolers' inferences about insides and causes. *Journal of Cognition and Development, 8,* 159-182.

Tyrrell, D.J., Stauffer, L.B., & Snowman, L.G (1991). Perception of abstract identity/difference relationships by infants. *Infant Behavior & Development, 14*, 125-129.

Wasserman, E.A., Fagot, F., & Young, M.E. (2001). Same-different conceptualizations by baboons (*Papio papio*): The role of entropy. *Journal of Comparative Psychology, 115*(1): 42-52.

Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology, 43,* 337–375.

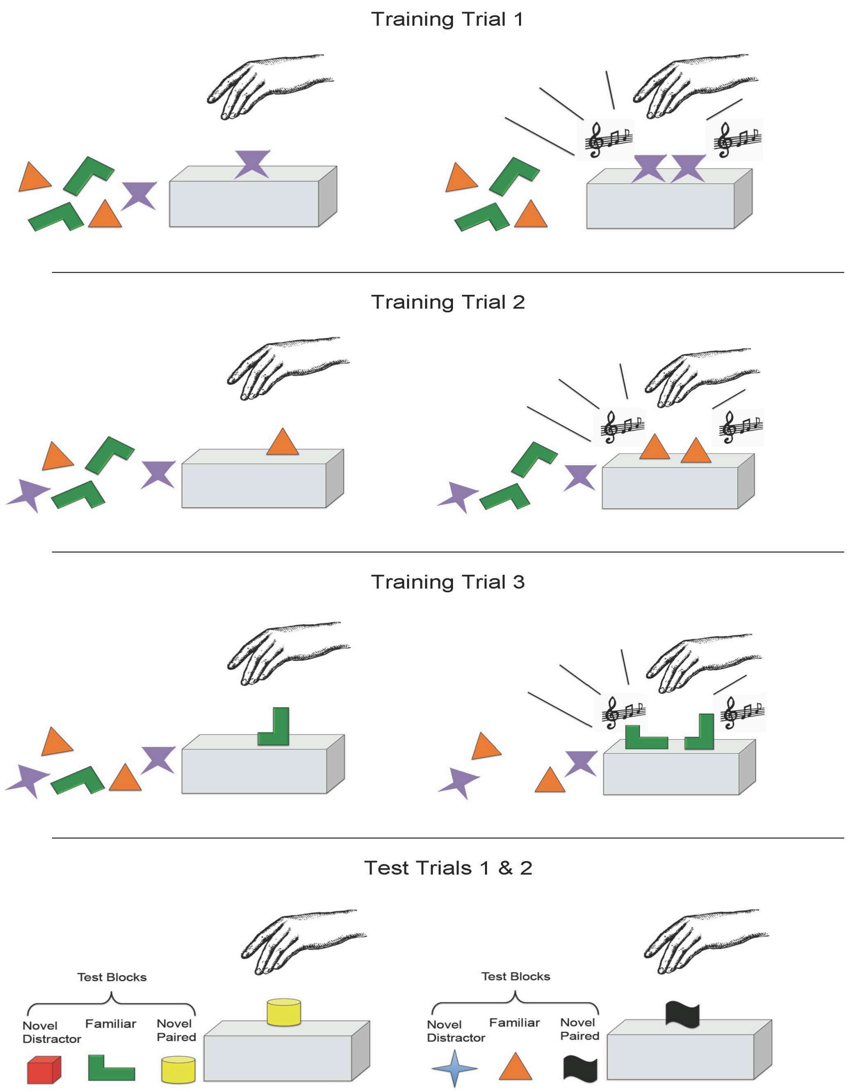

Fig. 1. Schematic representation of training and test trials in Experiment 1. On each training trial, a single block was placed on the toy (no activation) and then an identical block was added, activating the toy. This was repeated for all 3 training pairs. On each test trial, 3 test blocks (*novel paired block*, *familiar block*, *novel distractor block*) were presented. The experimenter then placed the target block on the toy, yielding no effect. The child was asked to select one test block to activate the toy.

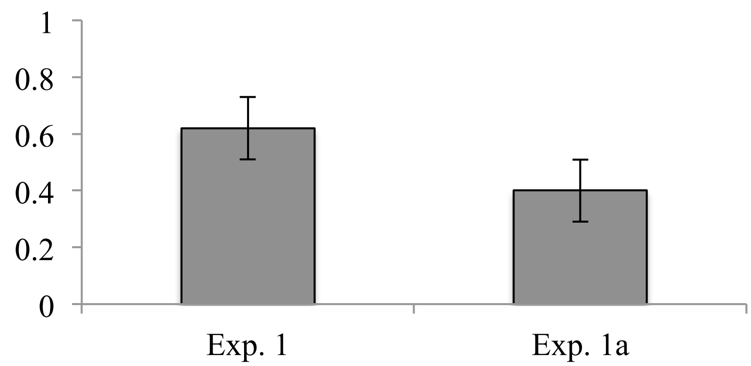

Figure 2. Percentage of 21-24 month olds in Experiement 1 and 1a who selected the matched pair (chance = .33).

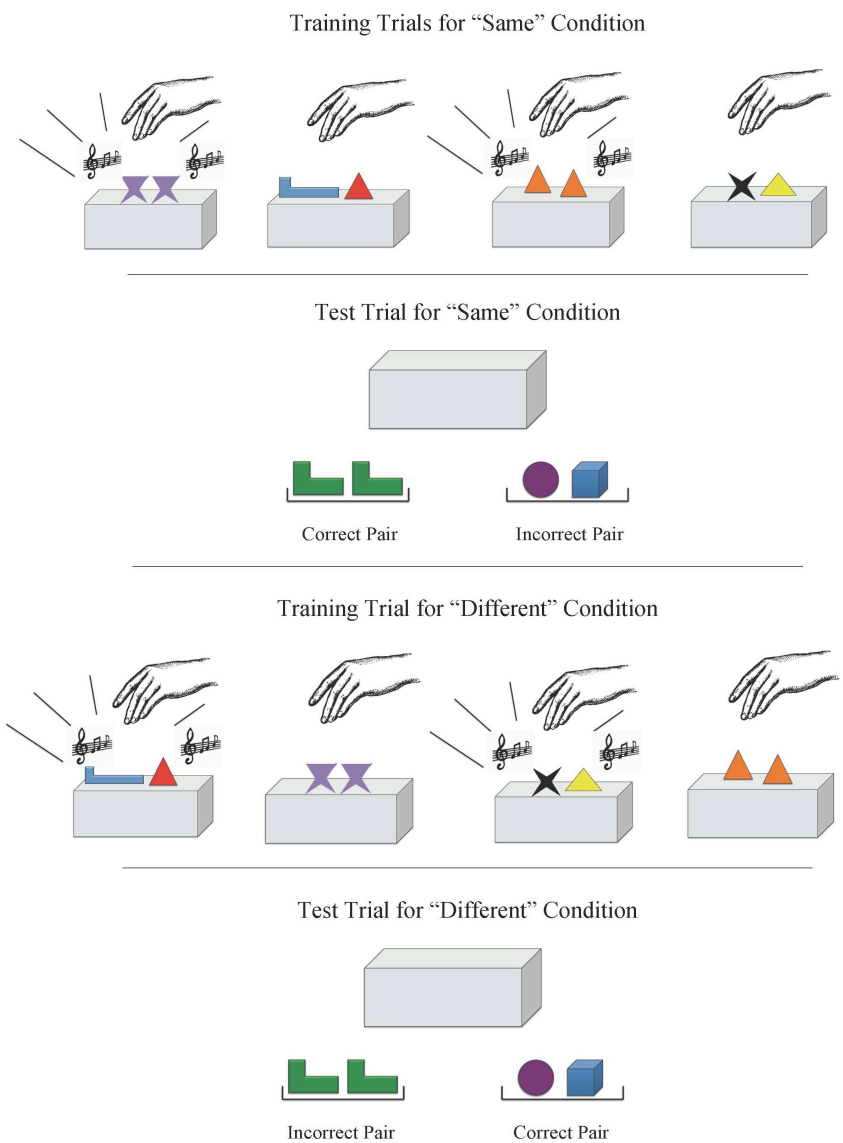

Fig. 3. Schematic representation of training and test trials in the *same* and *different* conditions of the relational match-to-sample task in Experiment 2. On each training trial, a pair of blocks were placed on the toy. In the *same* condition, the pairs of identical objects activated the machine. In the *different* condition, the pairs of distinct objects activated the machine. Participants observed four pairs (two causal and two inert). On each test trial, 2 pairs of test blocks were presented ("same" and "different"). The child was asked to select the pair that would activate the toy.

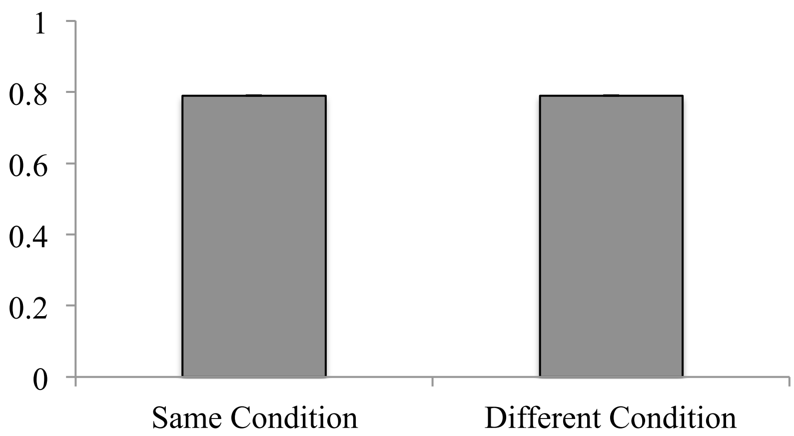

Figure 4. Percentage of toddlers in the *same* and *different* conditions in Experiment 2 who selected the correct pair during the test trial (chance = .50).