Preschool Children's Learning and Generalization of Continuous Causal Functions

Caiqin Zhou[1] (caiqin_zhou@brown.edu), **Rebekah A. Gelpí**[2,3] (rebekah.gelpi@mail.utoronto.ca),
Maria Iorini[1] (maria_iorini@brown.edu), **Christopher G. Lucas**[4] (clucas2@inf.ed.ac.uk),
& **Daphna Buchsbaum**[1] (daphna_buchsbaum@brown.edu)

[1]Department of Cognitive and Psychological Sciences, Brown University
[2]Department of Psychology, University of Toronto [3]Vector Institute
[4]School of Informatics, University of Edinburgh

Abstract

Many causal relations can be represented by continuous functions that map inputs to outputs. Can young children learn continuous causal functions and generalize them from observed data to new scenarios? We found that 4- and 5-year-olds can represent continuous functions with different abstract forms. After observing a few input-output pairs, children can accurately infer positive linear and step functions by predicting the outputs of novel input values. They also have emerging knowledge of negative linear and triangular functions. While children do not yet make consistently accurate predictions for these functions, they can distinguish these functions from the positive linear function and show inferential patterns that are consistent with the respective functions. Like adults and older children, preschoolers show an inductive bias towards positive linear functions. Their understanding of negative linear functions–which strongly requires inhibiting this inductive bias–improves with age.

Keywords: causal learning, function learning

Introduction

Children are surrounded by different forms of causal relationships between continuous variables in the world. For example, when sliding a dimmer switch, the further you move the slider to the "on" position, the brighter the light becomes. When building a sandcastle, adding more water to the sand initially helps it stick together better, until you have passed an optimal point, beyond which excess water starts reducing the structural integrity of the mixture. When using a loyalty punch pass at an ice cream shop, every additional point has no immediate effect, until you have reached a threshold that suddenly allows you to redeem the reward in full.

Functional relations are the relations mapping inputs to outputs and, as demonstrated in the above examples, we often use them to reason about the relation between cause and effect. To achieve desirable outcomes (e.g., building a stable sandcastle), we need to understand not only *that* a cause is related to an effect (e.g., water affects the structural integrity of the sand mixture), but also *how* manipulating the cause will change the effect. Understanding continuous causal relations in this general and abstract way allows us to make predictions about novel scenarios. Can young children learn different continuous relations between cause and effect, and generalize from limited experience to new examples?

While many causal relations are continuous, research on children's causal learning has largely focused on discrete causal rules. Many studies have used a *blicket detector* paradigm, in which a machine lights up and plays music when certain objects are placed on it (Gopnik & Sobel, 2000). After observing patterns of contingency between causes and effects (e.g., green blocks all activate the machine, but red ones do not), children as young as two years old can correctly infer and generalize the causal rules (e.g., activating the machine with a novel green block; Gopnik et al., 2001, 2004).

Recent work on causal relational reasoning has examined more abstract relations, showing that the representations of these relations emerge by preschool age. Goddu et al. (2020) showed children that a magic wand could transform an object from one state to another (e.g., turning a small apple into a big one). Three- and 4-year-olds correctly predicted how a new object would change. Similarly, after observing a shrinking machine and an enlarging machine, 2- and 3-year-olds chose the correct one to transform a misfitting object into the right size (Goddu et al., 2025). Notably, the causal framing enhanced children's learning of abstract relations compared to when the same stimuli were used in a non-causal context. This work suggests that children can represent discrete state changes. However, whether children's knowledge of abstract relations extends to continuous functions (e.g., the *magnitude* of size change in a magical transformation) is less clear.

Children have been shown to recognize the general forms of different continuous functions (e.g., monotonic versus U-shaped, linear versus sigmoid). Coates et al. (2023) showed children that lights and flowers differed in how their brightness or petal size changed over time (e.g., both quantities increased linearly, or increased and then decreased). When given multiple lights and flowers that changed based on different functions, 4- to 7-year-olds correctly matched the lights and flowers based on the common function that underlay each pair. This study shows that children can distinguish between different functional forms. However, recognizing that patterns match may not be the same as representing the underlying functions, which requires making predictions and generalizations beyond observed data points.

To our knowledge, school-age children are the youngest population in which the ability to generalize continuous functions has been studied. Zhou et al. (2024) presented 6- to 9-year-olds input-output pairs in a growth function (apple quantities in an orchard at several time points). Using these data, children successfully predicted outputs of novel input values (apple quantities at previously unobserved times). Notably, learning accuracy varied by functions–children inferred linear functions more accurately than Gaussian or exponential ones.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* ©2025 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Using a similar paradigm, Ebersbach and Wilkening (2007) have found that 7- to 13-year-olds predicted linear growth better than exponential growth. Why does performance differ across the functions? Do children have any inductive biases when reasoning about continuous functions?

Research on adult function learning provides strong evidence of an inductive bias in these experienced learners–they favor linear functions, especially those with a positive slope. These studies typically show learners input-output pairs, before asking them to predict outputs for new inputs (e.g., predicting a child's height from their age). Judgment accuracy for both previously observed and unobserved input values is higher for positive linear than negative linear functions, and for linear functions (regardless of directionality) than non-linear ones (e.g., U-shaped; Brehmer, 1974; Lucas et al., 2015). Adults also tend to extrapolate linearly–predicting points that fall on a straight line–even when the true relation is non-linear (DeLosh et al., 1997; Kalish et al., 2004).

An open question is where this inductive bias about continuous functions may have originated. If this bias emerges early, children should learn positive linear functions better than other functions. They might show a large performance gap between positive linear and other functions especially if they have limited attention or memory during learning. Alternatively, if adults' positive linearity bias is learned in school or from experience, such expectations might be weaker for children. These weaker expectations might help children revise their beliefs more readily than adults after observing belief-violating data (Lucas et al., 2014; Rich & Gureckis, 2018)–data that contradict the positive linear function in support of other functions. In other words, young children might learn non-linear functions more easily than older children and adults and perform similarly across different functions.

In the present study, we investigate whether (1) children can represent and generalize any continuous causal functions before school age. Specifically, can 4- and 5-year-olds learn different forms of continuous functions? (2) If so, do children exhibit an early emerging bias towards linear and positive-linear relations, similar to older children and adults?

Children interacted with machines that played music for different lengths of time when different-sized blocks were placed on them. Music durations varied as a function of block size (e.g., for the triangular function, medium-sized blocks activated the machine for a long time, while smaller and larger blocks activated it for short durations). Children observed the effects of a few different-sized blocks on the machine, before predicting the effect of block size on music length for blocks of unobserved sizes. We expected this causal framing to help children reason about relations between entities (Goddu et al., 2020; Walker & Gopnik, 2014).

We chose four functions: positive linear, negative linear, triangular, and step (Figure 1c). For each function, children answered the same set of prediction questions, which were designed to test if they understood key features of each function and distinguished between the functions. We hypothesized that preschoolers could learn about continuous functions, by performing above chance on the prediction questions for at least one of the functions. The functions also vary in abstract features (linearity and monotonicity), allowing us to test if children's learning is sensitive to these features.

Since linear functions are easier to learn than non-linear ones for older children (Ebersbach & Wilkening, 2007; Zhou et al., 2024), preschoolers might learn the two linear functions better than the non-linear ones. They might also learn the positive linear function more accurately than all the other functions (negative linear, triangle, and step), since adults are especially biased towards positive linear functions (Brehmer, 1974; Lucas et al., 2015). Children might also distinguish the functions by showing divergent patterns of predictions for the functions. However, if preschoolers do not yet share the (positive) linearity bias with older children and adults, they might show similar learning accuracy on all functions.

Methods

Design

Participants learned relations between the size of the blocks (X) and the length of time that the music machine activated (Y). Each participant encountered two functions. The first one was always a positive linear function; given that adults required the least training to learn positive linear functions, we expected that showing this function first would help familiarize children with the task[1]. The second function varied between participants and was either negative linear, triangular, or step. For each function, children answered the same set of test questions, which allowed us to compare children's response patterns across functions.

Participants

Our preregistered sample consisted of forty-eight 4- and 5-year-olds, eight of each year in each of the three function conditions ($M_{\text{age}} = 5.0$, range = 4.0–5.9 years; preregistration: https://osf.io/qku9h). Ten additional children were excluded due to parental/sibling interference (4), experimenter error (2), technical difficulties (2), or failure to complete the full session (2). Children participated in-person.

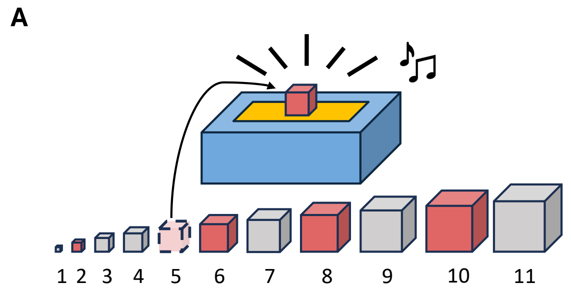
Materials

The music machines were two boxes, each with an iPad placed inside, directly below a transparent panel (Figure 1a). One experimenter interacted with children while another experimenter surreptitiously controlled the machine to activate (play music and light up) for different lengths of time, using a Bluetooth keyboard. We 3D-printed 11 different-sized white cubes, here referred to as Blocks 1–11, with the number corresponding to the block's side length in centimeters.

Procedure

Children completed two trials, each using one of the music machines. The experimenter introduced children to one of

[1] We discuss the impact of this fixed presentation order on p6.



B

Correct answer to each question for each function

	*"Will [**Block A**] make the machine go for longer, will [**Block B**] make the machine go for longer, or will they make it go for the **same** amount of time?"*			
	1 vs 2	**3 vs 9**	**7 vs 9**	**1 vs 11**
Pos linear	2	9	9	11
Neg linear	1	3	7	1
Triangular	2	Same	7	Same
Step	Same	9	Same	11

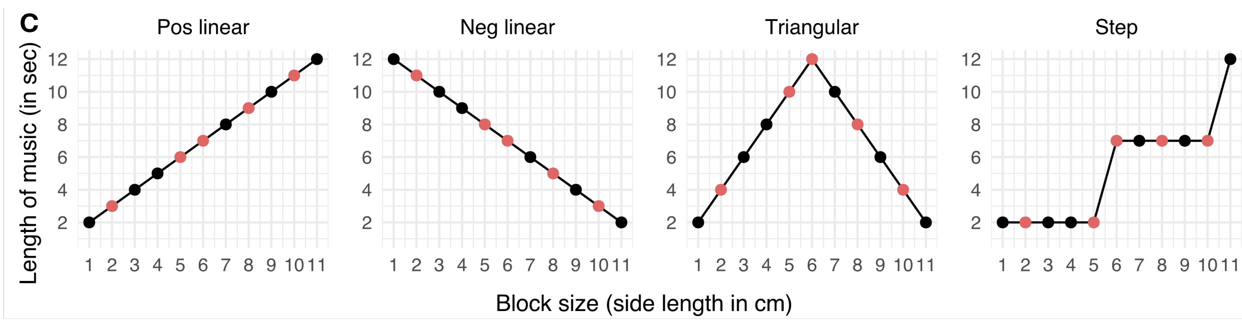


Figure 1: (A) Stimuli. All blocks were white, but the five blocks that children observed during learning were highlighted in red here for demonstration purposes. (B) Correct answers for each test question based on function. (C) Functions. Red dots represent the input-output pairs shown during learning.

the machines and the array of 11 blocks (ordered by size and spaced out evenly). The experimenter explained, "Different blocks make the machine play different amounts of music. Some blocks make it go for a short time, some blocks make it go for a long time, and some blocks make it go for some time in the middle. Your job is to figure out how long each of these blocks makes the machine play music."

In the learning phase, the experimenter asked children to try Blocks 2, 5, 6, 8, and 10 (fixed across participants) on the machine, one at a time in ascending order. This set of blocks was chosen to reveal key features of each function (Figure 1c) while allowing us to control for the size of the demonstration blocks across functions. For example, the five blocks produced increasingly longer music for the positive linear function; however, for the step function, Blocks 2 & 5 produced the same amount of music (2s), and Blocks 6, 8, & 10 produced the same amount of music (7s), revealing that inputs on the same interval produced the same effect and that the step occurred between Blocks 5 & 6. After children observed the effect of each block, the experimenter placed a visual reminder of the machine's duration of activation in front of that block. The reminders were paper strips with varying numbers of guitar icons, where the number of guitars (and hence the length of the strip) reflected the music duration in seconds.

In the test phase, children predicted the relative efficacy–in terms of music duration–of five pairs of different-sized blocks (Figure 1b). These questions were designed to test children's understanding of the key features of each function, as well as if they distinguished the functions by answering the same questions differently across the functions. For each pair, children answered a three-answer forced-choice question, "Will this block [Block A] make the machine go for longer, will this block [Block B] make the machine go for longer, or will they make it go for the same amount of time?" The first pair involved two previously observed blocks, 6 and 10, and served as a memory check (However, as preregistered, we did not exclude children based on their memory accuracy; we analyzed the data both with and without children who failed the memory check). The four test pairs all involved previously unobserved blocks: (Q1) 1 vs 2, (Q2) 3 vs 9, (Q3) 7 vs 9, (Q4) 1 vs 11. Whether the larger or smaller block in each pair was mentioned first was counterbalanced across participants. Children did not receive any feedback on their answers.

After children completed the first trial, the experimenter removed the first music machine and introduced the second machine of a different color. The experimenter told children, "Here is a different machine. This machine also plays music when we put the blocks on it, but the rule for how different blocks make this machine work is different. You need to figure out the new rule of how each block makes this new machine play music!" Children then went through the learning and test phases described above for the second trial.

Results

Main analysis: judgment accuracy

Our main research question was whether children could learn continuous functions, by performing above chance on the judgment questions for at least one of the functions. We also examined any inductive biases children might have–was their learning accuracy higher on the (positive) linear functions than the other functions? Lastly, was accuracy affected by the specific block pairs children were comparing (e.g., two blocks that were similar or highly different in sizes)?

To test these questions, we predicted children's judgment accuracy using function (positive linear, negative linear, triangular, step), question number (Q1: Blocks 1 vs 2; Q2: 3 vs 9; Q3: 7 vs 9; Q4: 1 vs 11), age (continuous), and the interaction between function and age. Judgment accuracy followed a binomial distribution (correct/incorrect) and each child answered multiple questions, so we used generalized logistic mixed models, with a random intercept for each child.

Main effect of function Learning accuracy varied significantly across functions (we used Type II Wald chi-square tests to assess the significance of fixed effects in our models; $\chi^2[3]$ = 49.88, $p < .001$). We further tested whether accuracy for each function was above chance ($\frac{1}{3}$; each forced-choice question had three answers: whether Block A or B would make more music, or if they would have the same effect), as well as how accuracy varied across functions (Figure 2). Performance was above chance on the positive linear function (estimated probability = .84, SE = 0.036, 95% CI = [.76, .90], $z = 8.78$, $p < .001$) and step function (estimated probability = .66, SE = 0.083, 95% CI = [.49, .80], $z = 3.69$, $p < .001$); performance was at chance on the negative linear function (estimated probability = .31, SE = 0.090, 95% CI = [.16, .51], z = -0.27, p = .79) and marginally below chance on the triangular function (estimated probability = .18, SE = 0.06, 95% CI = [.09, .33], z = -1.95, p = .051). Children more accurately learned the monotonically increasing positive linear and step functions, compared to the non-monotonic triangular function and the monotonically decreasing negative linear function. Children might have a bias towards inferring functions whose output values increase as input values increase.

Pairwise contrasts of judgment accuracy for the functions further confirmed children's superior performance on the positive linear and step functions. Specifically, the odds of correctly answering the test questions were higher for the positive linear function compared to the negative linear function (OR = 12.16, SE = 5.96, 95% CI = [4.66, 31.77], $z = 5.10$, $p < .001$), the triangular function (OR = 24.14, SE = 11.24, 95% CI = [9.70, 60.10], $z = 6.84$, $p < .001$), and the step function (OR = 2.76, SE = 1.05, 95% CI = [1.31, 5.82], $z = 2.67$, $p = .008$). Notably, even between the two increasing functions, children still predicted the linear function more accurately than the step function; children might prefer functions that increase, particularly those that increase linearly. Furthermore, the odds of correctly answering the questions were also higher for the step function compared to the negative linear function (OR = 4.40, SE = 2.53, 95% CI = [1.43, 13.56], $z = 2.58$, $p = .01$) or triangular function (OR = 8.74, SE = 4.83, 95% CI = [2.96, 25.80], $z = 3.93$, $p < .001$). Accuracy on the negative linear and triangular functions did not differ (OR = 1.99, SE = 1.13, 95% CI = [0.65, 6.08], $z = 1.20$, $p = .23$), suggesting that children did not distinguish between the two functions that did not increase monotonically.

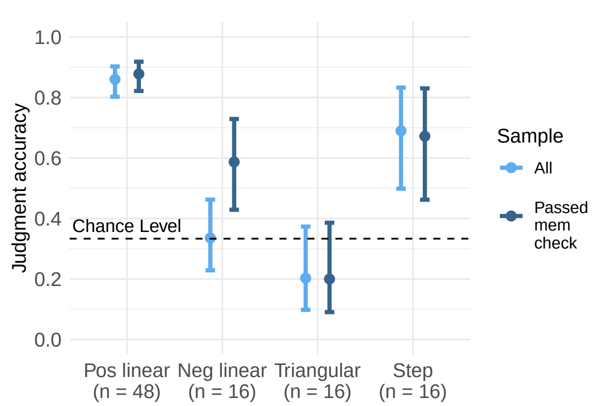


Figure 2: Judgment accuracy by function and memory accuracy, with 95% CIs.

Main effect of question Learning accuracy also varied across test questions ($\chi^2[3] = 8.11$, $p = .044$). In other words, the specific blocks for comparison (e.g., blocks that were similar or highly different in sizes) affected the ease of judging the relative efficacy of the blocks. Pairwise contrasts showed that the odds of correctly answering Q4 (Blocks 1 vs 11, smallest vs largest) were significantly higher compared to Q1 (Blocks 1 vs 2; OR = 2.79, SE = 1.06, 95% CI = [1.33, 5.88], $z = 2.70$, $p = .007$) and Q3 (Blocks 7 vs 9; OR = 2.15, SE = 0.82, 95% CI = [1.02, 4.52], $z = 2.02$, $p = .043$). The comparisons of other pairs of questions did not reach significance (all $|z| \leq 1.62$, all $p > .10$.) Children might be better at comparing output values of highly different inputs (e.g., Blocks 1 vs 11) than similar inputs (e.g., Blocks 1 vs 2). Put differently, children might be more sensitive to the overall trend of the function (e.g., whether it was increasing or decreasing) than its specific shape within a narrow region (e.g., whether two adjacent blocks would have the same or different effects).

Interaction effect of age and function The main effect of age did not reach significance ($\chi^2[1] = 2.27$, $p = .13$), however, the interaction between age and function was significant ($\chi^2[3] = 10.34$, $p = .016$). We further tested how the age effect varied across functions (Figure 3). Age only improved accuracy on the negative linear function (b = 0.23, SE = 0.070, 95% CI = [0.090, 0.37], $z = 3.24$, $p = .001$) but not the others (all $|z| \leq 1.59$, all $p > .11$). Children might have shown the largest improvement on the negative linear function because it requires fully inhibiting the positive linearity bias; children have to suppress the urge to respond based on a positive linear pattern and have to reverse it to learn the opposite relation. Inhibitory control and cognitive flexibility develop rapidly during the preschool years (Zelazo et al., 2003), which might explain the drastic age effect here.

Memory and performance As preregistered, we also conducted the main analysis with only children who passed the memory check (Blocks 6 vs 10, two previously observed

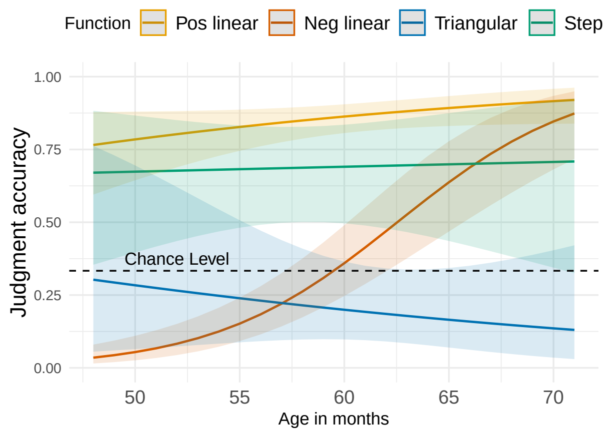


Figure 3: Judgment accuracy across age, with a separate logistic regression for each function and 95% CIs.

blocks) in both trials (n = 32 out of 48). These children performed above chance on the negative linear function (estimated probability = .58, SE = 0.11, 95% CI = [.36, .78], $z = 2.19$, $p = .029$; Figure 2). Their accuracy on the other functions and the age effect were consistent with the results based on the full sample. These children who had high task engagement and accurate memory during learning performed better on the more challenging negative linear function, which required overcoming a possible positive linearity bias.

Secondary analysis: response pattern

We also investigated if children differentiated the functions, by responding to the test questions differently across the functions (Figure 4). For example, if children understood the decreasing nature of the negative linear function, they should predict that the smaller block (out of any pair of blocks) would activate the machine for longer. Compared to when learning the positive linear or step functions, children should pick the smaller block more often when learning the negative linear and triangular functions, which both have a negative linear component. We could not run a categorical model to analyze the frequency of picking the three answers across functions because some answers were never chosen for certain function and question combinations. Therefore, we created three binary variables, one for each answer, indicating whether children chose the smaller block, the larger block, or ”same,” respectively (e.g., choosing ”smaller” would correspond to the values of 1, 0, and 0). We ran three generalized logistic mixed effects models, one with each binary variable as the outcome. In each model, we predicted whether children chose that answer using function and question number, treating the positive linear function as the reference category.

The probability of choosing the smaller block indeed varied across functions ($\chi^2[3]$ = 42.03, $p < .001$). Consistent with the correct pattern of responding (Figure 4), the odds of choosing the smaller block were higher for the negative linear function compared to the positive linear function (OR = 21.38, SE = 11.98, 95% CI = [7.13, 64.13], $z = 5.47$, $p <$ .001) and higher for the triangular function (which has a decreasing component) compared to the positive linear function (OR = 10.71, SE = 5.46, 95% CI = [3.94, 29.08], $z = 4.65$, $p < .001$). As expected, the odds of choosing ”smaller” did not differ for the step and positive linear functions (OR = 0.58, SE = 0.49, 95% CI = [0.11, 3.09], $z = -0.64$, $p = .52$).

For the step function, if children knew that inputs on the same interval lead to the same output, they should pick ”same” more often when learning this function compared to the positive linear function. The probability of picking “same” varied marginally by functions ($\chi^2[3]$ = 7.04, $p =$ .071). As expected, the odds of picking “same” were higher for the step function than the positive linear function (OR = 2.85, SE = 1.36, 95% CI = [1.12, 7.28], $z = 2.19$, $p = .029$).

Lastly, children should choose ”larger” most often for the positive linear function, for which it is the correct answer to every question. The probability of choosing ”larger” differed across functions ($\chi^2[3]$ = 40.7, $p < .001$). As expected, the odds of choosing the larger block were higher for the positive linear function compared to the negative linear (OR = 16.32, SE = 7.98, 95% CI = [6.26, 42.57], $z = 5.71$, $p < .001$), triangular (OR = 5.23, SE = 2.25, 95% CI = [2.25, 12.15], $z = 3.84$, $p < .001$), and the step (OR = 2.30, SE = 0.95, 95% CI = [1.02, 5.17], $z = 2.01$, $p = .044$) functions.

Discussion

Our results are the first to suggest that 4- and 5-year-olds already have some knowledge of continuous functional relations between cause and effect. They accurately infer the positive linear and step functions from observing a few data points and rely on these relations to predict the outputs of novel input values. Like adults and older children, preschoolers may be biased towards inferring linear or monotonically increasing functions. Children do not yet perform above chance on the negative linear or triangular functions; however, they show different response patterns that align with the respective functions and that reflect a differentiation of those functions from the positive linear function.

Why do children struggle with the negative linear function, especially when they perform well on the non-linear step function? While past work has shown that children learn linear functions more accurately than non-linear ones (Ebersbach & Wilkening, 2007; Zhou et al., 2024), neither study has explicitly tested increasing linear functions against decreasing ones. Moreover, adults show superior performance on positive linear than negative linear functions (Brehmer, 1974; Lucas et al., 2015). One possibility is that children are biased towards learning both linear functions (regardless of directionality) and increasing functions (regardless of functional form, as suggested by their accuracy of forecasting exponential growth versus decline; Ebersbach et al., 2008). When learning functions that involve both features, the positivity bias may outweigh the linearity bias. Another possibility is that children are only biased towards *positive* linearity.

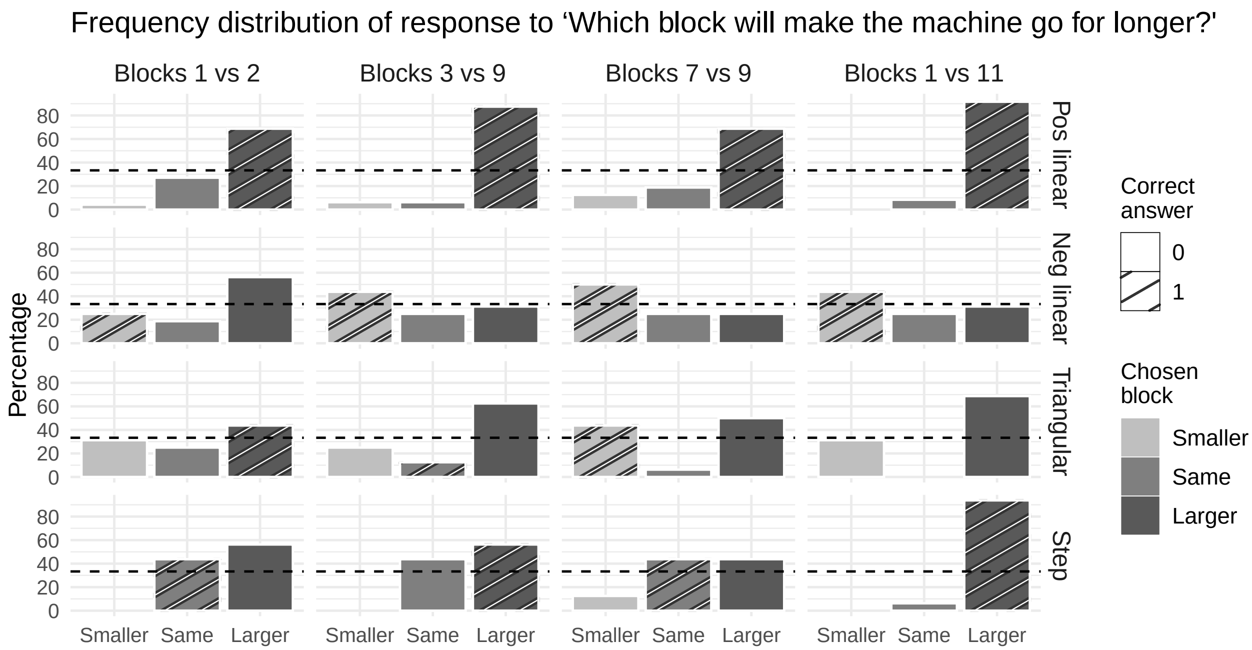


Figure 4: Frequency distribution of responses to each test question for each function. Dashed lines indicate chance level (1/3).

If children indeed have a positive linearity bias, they may have knowledge of negative linear (and triangular) functions but fail to demonstrate it due to the need to inhibit the bias. For the negative linear function, children always need to pick answers that are the opposite of what is true for the positive linear function. Children improve drastically on the negative linear function in the 4-6 year window we test here, when inhibitory control is known to develop rapidly (Zelazo et al., 2003). Providing anecdotal support for this explanation, some children spontaneously verbalize the negative linear relation, only to answer the questions based on the positive linear pattern; other children answer the first two questions based on the negative linear relation, before switching back to the positive linear pattern for the remaining questions.

Even if children do not perform above chance on the negative linear and triangular functions yet, they do show some knowledge of the respective functional forms. They correctly pick the smaller block more often for the negative linear and triangular functions (which both have a decreasing component) compared to the increasing functions. This aligns with findings that older children correctly infer the shape of nonlinear functions even if accuracy is higher on linear functions (Zhou et al., 2024). The triangular (and step) functions may also be more complex; children may have some general ideas (e.g., the former is non-monotonic and the latter has intervals) but not all the details (e.g., where the peak or step occurs).

We have chosen to always present the positive linear function first to facilitate task comprehension, as adults require the least training to learn this function; however, this fixed presentation order could have differentially impacted the learning of the subsequent functions. Specifically, the positive linear function shared 0%, 25%, and 50% of the answers with the negative linear, triangular, and step functions, respectively (Figure 1b). Although the triangular function shared more answers with the positive linear function than the negative linear function did, children who passed the memory check performed at chance on the former but above chance on the latter. Furthermore, children chose "same" more often for the step function compared to the positive linear function, both when "same" was the correct answer and when "larger", the answer under the positive linear function, was the correct answer (Figure 4). These results partially challenge the interpretation that performance on the second function reflects its similarity to the first function.

Another concern is that children could have interpreted the X axis as reflecting the block's volume instead of side length, perceiving the axis non-linearly (a 1 cm change of side length could lead to different changes in volume). To facilitate accurate perception of both axes, we spaced out the blocks evenly and provided visual reminders of music duration (paper strips of varying lengths). Non-linear perception does not seem to explain the response patterns (e.g., children often fail on the negative linear function by picking "larger" based on the positive linear function, not by picking "same" from an inability to discriminate between the smaller blocks).

Using an innovative task, we found that preschoolers can learn continuous causal functions and may have a positive linearity bias. They do not yet accurately infer non-monotonic or decreasing functions, which may be due to their strong inductive biases, poor inhibitory control, partial knowledge of the functions, or difficulty with representing the functions. We hope to disentangle these possibilities in our future work.

Acknowledgments

We thank members of the Computational Cognitive Development Lab for their valuable feedback on this study and their help with recruitment and data collection. We also thank the participating research sites and families. This work was supported by a Brown University OVPR Seed Grant.

References

Brehmer, B. (1974). Hypotheses about relations between scaled variables in the learning of probabilistic inference tasks. *Organizational Behavior and Human Performance*, *11*(1), 1–27.

Coates, N., Siegel, M., Tenenbaum, J., & Schulz, L. (2023). Representations of abstract relations in early childhood. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *45*(45).

DeLosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(4), 968.

Ebersbach, M., Lehner, M., Resing, W. C., & Wilkening, F. (2008). Forecasting exponential growth and exponential decline: Similarities and differences. *Acta Psychologica*, *127*(2), 247–257.

Ebersbach, M., & Wilkening, F. (2007). Children's intuitive mathematics: The development of knowledge about nonlinear growth. *Child Development*, *78*(1), 296–308.

Goddu, M. K., Lombrozo, T., & Gopnik, A. (2020). Transformations and transfer: Preschool children understand abstract relations and reason analogically in a causal task. *Child development*, *91*(6), 1898–1915.

Goddu, M. K., Yiu, E., & Gopnik, A. (2025). Causal relational problem solving in toddlers. *Cognition*, *254*, 105959.

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.

Kalish, M. L., Lewandowsky, S., & Kruschke, J. K. (2004). Population of linear experts: Knowledge partitioning and function learning. *Psychological Review*, *111*(4), 1072.

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.

Lucas, C. G., Griffiths, T. L., Williams, J. J., & Kalish, M. L. (2015). A rational model of function learning. *Psychonomic bulletin & review*, *22*(5), 1193–1215.

Rich, A. S., & Gureckis, T. M. (2018). The limits of learning: Exploration, generalization, and the development of learning traps. *Journal of Experimental Psychology: General*, *147*(11), 1553.

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

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., Chiang, J. K., Hongwanishkul, D., Schuster, B. V., Sutherland, A., et al. (2003). The development of executive function in early childhood. *Monographs of the society for research in child development*, i–151.

Zhou, C., Gelpi, R., Lucas, C., & Buchsbaum, D. (2024). Can children learn functional relations through active information sampling? *Proceedings of the Annual Meeting of the Cognitive Science Society*, *46*.