Cognition 254 (2025) 105959

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Causal relational problem solving in toddlers

Mariel K. Goddu [a,b,c,d], Eunice Yiu [a,*], Alison Gopnik [a]

[a] Department of Psychology, University of California – Berkeley, Berkeley 94704, USA
[b] Department of Philosophy, Stanford University, Palo Alto 94305, USA
[c] Institut für Philosophie, Freie Universität Berlin, Berlin 14195, Germany
[d] Centre for Advanced Study in the Humanities: Human Abilities, Berlin 10969, Germany

ARTICLE INFO

Keywords:
Problem-solving
Relational reasoning
Causal reasoning
Human development

ABSTRACT

We investigate young children's capacity for "causal relational reasoning": the ability to use relational reasoning to design novel interventions and bring about novel outcomes. In two experiments, we show that 24–30-month-old toddlers and three-year-old preschoolers use relational reasoning in a causal problem-solving task. Even toddlers rapidly inferred relational causal rules and applied this knowledge to solve novel problems—thus demonstrating both surprisingly early competence in relational reasoning and sophisticated causal inference. In both experiments, children observed a handful of trials in which a mechanistically opaque machine made objects larger or smaller. When prompted to solve a new problem, they used the machine to change the relative size of a novel object – even though its appearance and absolute size differed from previous observations, and even though subjects had never seen the machine generate objects of the required size before. This suggests that children quickly inferred abstract causal relations and then generalized these relations to determine which intervention would bring about the novel outcome required to solve the problem. These findings suggest a close link between early relational reasoning and active causal learning and inference.

1. Introduction

Everyday action hinges on understanding and influencing causal relations: by acting, we change the world to bring about the outcomes we desire. We leverage our causal knowledge whenever we intervene to make things happen, such as bouncing a ball by dropping it, catching someone's attention by waving our hand, ringing a bell by pulling a cord, cooling a drink by adding ice, or making a campfire bigger by fanning it. Even infants learn and apply direct causal relations to influence their social and physical environments. In the first months of life, babies learn to cry in targeted ways to attract caregivers' attention and to coo, smile, and imitate facial expressions to prolong social interactions. Gradually, they also learn contingencies between gross motor movements and changes in the physical environment, such as swatting at objects to move them or shaking a rattle to hear a sound (Buchanan & Sobel, 2011; Bullock, 1984; Goddu & Gopnik, 2024; Sobel & Legare, 2014).

One influential framework for understanding causal reasoning in humans and other animals is 'interventionism' (Pearl, 2000; Pearl, 2009; Woodward, 2005). On the interventionist view, causal relations are characterized as relations between variables: a variable C is a cause of another variable E if changing the value of C changes the value of E. Because of this, the interventionist view is sometimes described in terms of "difference-making": a cause is something that "makes a difference" to something else.

Some instances of causal reasoning involve relatively non-specific expectations about the kinds of differences that inverventions will make. It is sometimes possible to predict that an action will lead to *some* change, without knowing precisely what the change will be. (Consider interventions on stochastic and complex systems—such as a child blowing on a dandelion in a blustery wind, or a dogwalker pulling on leashes to steer seven dogs around a corner.) In many other cases, however, we have more precise expectations about how our interventions will change the world. For example, pulling a cord to signal an emergency involves knowing how to make a difference to whether the alarm is ringing. Children as young as 16 months can learn that particular properties of objects are difference-making for specific effects in this way. For example, they can learn that red blocks—but not yellow or blue blocks—cause a machine to light up and play music (Goddu & Gopnik, 2024).

In other situations, we seem to have even more fine-grained expectations about the way that our actions will change the world. Adults

* Corresponding author.
E-mail addresses: goddu@stanford.edu (M.K. Goddu), ey242@berkeley.edu (E. Yiu), gopnik@berkeley.edu (A. Gopnik).

https://doi.org/10.1016/j.cognition.2024.105959
Received 24 January 2024; Received in revised form 7 September 2024; Accepted 9 September 2024
Available online 27 September 2024
0010-0277/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

often understand not only *that* their actions will make a difference, but also the *kind of difference* they will make. Stoking a campfire, for example, involves not only poking it to simply change it—since unskillful poking can easily result in the fire's going out. Rather, successfully stoking a campfire involves making specific, directional differences to the fire's size and temperature: that is, poking it in such a way to make the sticks *closer,* so that the fire becomes *bigger* and *hotter*. Causal knowledge like this is especially useful for generalization to new situations. For example, your knowledge how to make a campfire bigger by increasing the closeness of sticks may enable you to make similar interventions to the logs in a bonfire or the coals of a barbecue. Existing studies have yet to investigate children's causal learning in contexts that involve learning more abstract relations between events—such as relations in which a *larger* block leads to a *brighter* light, or *louder* music. This is the starting point for the present investigation.

How is causal understanding related to relational reasoning? We begin with the observation that 'difference' (in "difference-making") is an *abstract relation.* A relation is a notion that applies to, or holds between, multiple entities—it is not a property that a single thing has on its own. The capacity to appreciate abstract relations—notions like 'same,' 'different,' 'opposite of,' 'bigger than,' 'smaller than,' and 'in between'—forms the basis of *relational reasoning.*

Relational reasoning is a powerful cognitive capacity that is thought to underpin the abilities to reason analogically, understand figurative language, and solve "insight" tasks, among other skills (Gentner, 1983; Holyoak & Thagard, 1994; Sternberg, 1977). A large literature on the development of relational reasoning has suggested that it is relatively late-emerging and depends on social factors to develop. Such factors include the acquisition of relational language (e.g., words like "big" and "small") and sociolinguistic cues (e.g., an adult's explicit prompts to compare multiple exemplars) (Gentner, 2003; Goswami, 2013; Richland, Morrison, & Holyoak, 2006).

However, more recent studies (see below) show that even young children can appreciate abstract relations in causal contexts—in particular, when they are operationalized as the beginning and ending states of causal transformations (Goddu, Lombrozo, & Gopnik, 2020). Could children use such knowledge to design novel causal interventions? Could they learn and apply abstract causal relations to solve a new problem in a new way?

On the one hand, the capacity to learn and generalize abstract causal difference-making relations seems like an adaptive, action-relevant skill: agents who have it would be better at designing new interventions in the face of novel challenges (Adolph & Hoch, 2019). In particular, it could be an important source of generalization. On the other hand, a large literature on the development of relational reasoning suggests that young children often struggle to learn and attend to abstract relations, and so might have difficulty using those relations to make new causal interventions or solve causal problems. We outline the literature supporting each of these competing hypotheses in more detail below.

1.1. Development of relational reasoning

Relational reasoning refers to the ability to reason using *abstract, structural* similarities (i.e., formal, functional, or other logical similarities) between entities—e.g., 'opposite of'; 'inside of'; 'larger than'—instead of *concrete, featural*, or *thematic* similarities (e.g., 'red'; 'fruit'; 'found in the kitchen'). Many tasks that have been used to study relational reasoning involve matching games with visual stimuli printed on flashcards. For example, a subject might see a flashcard (the "sample") with two identical strawberries. If they choose to match the sample with another card showing a strawberry and an apple, their match reflects a relatively superficial similarity (e.g., matching 'fruit' with 'fruit', or matching 'red' with 'red'). By contrast, if they instead choose to match the sample card depicting two identical strawberries to a card depicting two identical dinosaurs, their match is based on relational similarities (i.e., matching 'same' with 'same'). Other tasks require subjects to complete sentences or analogies. For example, a subject might be prompted to choose between FOOT (relational choice) and SHOE (thematic choice) to complete "GLOVE:HAND :: SOCK:??".

In these matching games, preschool-aged children tend to choose "object matches" based on superficial attributes rather than "relational matches" based on structural commonalities like spatial arrangement, function, or higher-order categories (Gentner, Anggoro, & Klibanoff, 2011; Gentner & Namy, 1999; Goddu, Lombrozo and Gopnik, 2020; Hammer, Diesendruck, Weinshall, & Hochstein, 2009). In the examples provided above, this means that they typically prefer to match "strawberry-strawberry" with "strawberry-apple" instead of "dinosaur-dinosaur", and they select SHOE over FOOT to complete GLOVE:HAND:: SOCK:??. Authors have frequently interpreted failures on these tasks as evidence that young children struggle to appreciate abstract relations—or even that they cannot reason relationally at all (e.g., Christie & Gentner, 2010; Christie & Gentner, 2014).

However, several recent studies have challenged this longstanding assumption. Instead of measuring relational reasoning by using matching tasks with static stimuli, this new work suggests that young children can attend to, learn from, and understand abstract relations in *causal* contexts. One study demonstrated that 4- to 6-year-olds intuitively mapped the abstract forms of effects to their causes: for example, they thought that a discretely varying effect (e.g., a beeping tone) was more likely to be explained by a discretely varying cause (e.g., pressing a button) rather than by a continuously varying one (e.g., turning a knob) (Magid, Sheskin, & Schulz, 2015). Another study directly tested relational learning and generalization in the context of novel causal transformations. Three-year-olds saw animations in which a wizard first made an apple larger when she waved her wand. Next, she waved her wand and made a dog larger. When the children were then asked to make a prediction about the effect of the agent's next action—"What do you think she's going to do next?"—they were more likely to choose an animation in which a novel dice grew larger (relational match—one that exhibited *the same kind of difference*) over one in which the previously seen apple became flattened (perceptual match) (Goddu, Lombrozo and Gopnik, 2020). Still other studies have shown that certain higher-order relations, such as '*same-different*' relations, are already accessible in toddlerhood and possibly even infancy (Dewar & Xu, 2010; Ferry, Hespos, & Gentner, 2015; Hochmann, Mody, & Carey, 2016; Walker & Gopnik, 2014). These findings challenge the traditional notion that the development of relational reasoning depends on 'sociolinguistic' cues, such as relational language or prompts to compare exemplars (Christie & Gentner, 2010; Christie & Gentner, 2014). They also suggest that early relational reasoning may emerge in causal contexts before it becomes apparent in others.

1.2. Development of causal reasoning

From an 'interventionist' perspective, the criterion for genuine causal understanding is that a reasoner can intervene on the cause to bring about the effect (Gopnik & Wellman, 2012; Pearl, 2000; Pearl, 2009; Woodward, 2005). This distinguishes causal understanding, as in the cases we described in the beginning (*making* the flames bigger, *making* the bell ring), from the simpler ability to predict one event after observing another—the ability that underlies simple associative learning and pattern-matching.

According to this 'causal intervention' criterion, even infants can understand and learn some basic causal relations. For example, with their foot tethered to a mobile suspended above their crib, 3- to 4-month-olds can learn that their kicking causes the mobile to move (Piaget & Cook, 1952; Sloan, Jones, & Kelso, 2023). Even younger children may cry to bring about targeted changes in the social world (e.g., to summon their caregivers; Goddu and Gopnik, 2024). By eight months, children seem to appreciate bidirectional relations between action and effect, as in shaking a rattle (Paulus, Hunnius, Van Elk, & Bekkering, 2012). And by 12 months, children can already interact purposefully with touchscreen devices (Ahearne, Dilworth, Rollings, Livingstone, & Murray, 2016). Eighteen-month-olds can use "intuitive physics" knowledge (Carey & Spelke, 1996; Gilhooly & Murphy, 2005; Spelke, 2014; Spelke & Kinzler, 2007; Ullman, Spelke, Battaglia, & Tenenbaum, 2017) to solve "insight" and "means-ends reasoning" tasks, which require manipulating objects in novel ways. For example, they can apply their knowledge of contact and support relations to use a new stick to retrieve a distant toy without a period of trial and error (Uzgiris & Hunt, 1975; Willatts, 2013).

All of these examples involve "first-personal" causal interventions that occur in the context of children's own goal-directed activity (Goddu & Gopnik, 2024). But numerous studies indicate that young children are also capable of even more sophisticated causal reasoning, such as trying a new cause to bring about an outcome that they have only observed someone else perform. For example, a large literature using "blicket detector" paradigms suggests that when two-year-olds observe statistical contingencies between an experimenter's actions and outcomes (e.g., seeing that placing a red cube on a machine is followed by music), they infer causal rules and generalize them to novel instances (e.g., place a new red triangle on the machine to cause music) (Gopnik et al., 2004; Gopnik, Sobel, Schulz, & Glymour, 2001; Gopnik & Sobel, 2000). One study also demonstrated that 3-year-olds learned which of several kinds of causal variables were "difference-making" for an effect in novel biological and mechanical systems. After observing interventions on variables in the system, participants then successfully generalized: when confronted with novel values of the variables and asked to produce an effect they had never seen, they chose to intervene on the relevant variable—i.e., the one that was previously observed to be 'difference-making' (Goddu & Gopnik, 2020).

However, all of these experiments involve relatively simple "difference-making" relations between particular features of objects and events (e.g., a block makes the machine light up or causes music to play). Existing studies have not investigated children's causal reasoning and generalization in tasks that require learning about more abstract relations between events—such as relations in which a *larger* block leads to a *brighter* light, or *louder* music. In addition, the handful of studies that show positive evidence for early relational reasoning in causal contexts (see previous section) have only tested children's ability to make *predictions*, rather than *interventions* (Dewar & Xu, 2010; Ferry et al., 2015; Goddu et al., 2020; Hochmann et al., 2016; Walker & Gopnik, 2014). To show that children genuinely understood these relations in a causal way, we would need to show that they could use the relation to intervene appropriately and bring about a novel outcome. Given the apparent difficulty that young children have in relational reasoning, it is possible that this may be too complex for them to manage.

1.3. The present study: Causal relational reasoning

The present study tests whether young children can perform "causal relational reasoning" – which we define as the ability to generalize abstract relations to make novel causal interventions. The tasks require that participants attend to the form of "differences made" to several objects during a handful of observation trials, and then generalize these relations to a new, perceptually dissimilar object—that is, apply them to solve a novel problem. Our tasks thus test for both early relational reasoning as well as the ability to select an appropriate causal intervention to produce a novel outcome (a hallmark of causal understanding).

Crucially, in contrast to shaking a rattle or moving a mobile, the stimuli that we use in the present experiments cannot be understood by using "intuitive physics," the capacity to reason about the dynamics and interactions of ordinary physical objects (Carey & Spelke, 1996; Gilhooly & Murphy, 2005; Spelke, 2014; Spelke & Kinzler, 2007; Ullman et al., 2017). The purpose of this design is two-fold. First, it ensures that children's learning about the events is limited to relational information. Second, it tracks a real-world phenomenon: adults possess the capacity for causal learning even in scenarios where the underlying process or mechanism is hidden or unknown, such as in many modern electronic and digital technologies. For instance, adjusting a thermostat, a dimmer switch, or a brightness slider on a smartphone are actions that most people perform quite automatically, despite having little to no understanding of the underlying mechanisms (Gärdenfors & Lombard, 2020). When adults reason about mechanistically opaque causal relations to use these technologies, they seem to grasp something about their abstract *form*: that is, they know the *general way* that changing the cause variable will make differences to the effect variable. For example, when trying out a new electric keyboard at the music store, we seem to learn more than particular, point-like relations such as, "turning the knob from 2 to 3 *makes a difference*" to the volume. Rather, we learn that "turning it *this way* makes it *louder.*" This abstract causal knowledge enables us to predict the outcomes of a wide range of interventions that have not yet been attempted (e.g., predicting what will happen if we turn it to 11). It therefore allows much wider causal generalizations than a simple understanding of causal properties, a particularly important capacity for understanding new tools like sliders or knobs. If very young children exhibit the same kind of learning, this might provide a clue about how observational causal learning (i.e., learning by observing others' actions; Meltzoff, 1995; Want & Harris, 2002) can be such an effective way to transmit important and generalizable skills (Csibra & Gergely, 2009; Tomasello, 2009).

In Experiment 1, we demonstrate that three-year-olds learn the abstract forms of novel difference-making relations caused by two novel "machines" from a small number of examples. They then generalize their relational understanding to select an appropriate intervention to produce a novel outcome in the service of solving a new problem. Experiment 2 replicates these findings with 24- to 30-month-olds. The findings suggest that children are capable of "causal relational reasoning": combining relational and causal thinking to recognize novel interventions and solve problems.

2. Materials and methods

Experiment 1 was conducted from 2016 to 2017. Experiment 2, collected from 2019 to 2020, had its procedures and analysis preregistered on https://aspredicted.org/PQK_2MV (anonymized PDF).

2.1. Participants

136 participants were recruited across the two experiments. In Experiment 1, participants were 36 preschoolers (M_{age} = 41.3 months, SD_{age} = 4.37 months, range = 36–48 months, 22 females), with three excluded for parental interference (2) and failure to provide a response (1). Experiment 2 had 100 toddlers (M_{age} = 27.3 months, SD_{age} = 2.7 months, range = 24–30 months, 52 females), with 20 excluded for parental interference (10), experimenter error (4), failure to provide a response (3), fussiness (2), or interruption (1). Data with "parental interference" universally involved utterances of relational language (e.g., "Look! It got smaller!") and were excluded to rule out the possibility that verbal labeling of abstract relations was driving causal relational learning.

2.2. Materials

Materials for Experiment 1 included two 25 cm^3 silver boxes, one with an 8 cm × 10 cm red felt flap and the other with a matching purple felt flap. During the demonstration phase, one 4-cm-wide red ball and one 4-cm-wide blue sponge were used. One box contained a larger version of the pair (both 8-cm-wide) and the other contained a smaller version (2-cm-wide). Half of the participants saw that the box with the red flap made objects smaller, while another half saw that it made them larger. Other materials included a 7.5-cm-tall plastic monkey figurine, a 5-cm-wide blue crown and 5-cm-wide green beanie that fit on the monkey's head, an undersized 2.5-cm-wide yellow paper hat, and an oversized 10-cm-wide one; see Fig. 1. A twinkle sound effect from an iPhone accompanied each object's (unobserved) "transformation".

In Experiment 2, a large black box (50 × 29.5 × 26 cm) with a yellow and blue door on each side housed a metal platform on a rotating stage. The setup allowed the experimenter to magnetically attach and rotate away objects behind the apparatus, replacing them with a differently sized version. This facilitated the illusion of a "change machine" that transformed object sizes. A 4-cm-wide spherical red character with googly eyes and a magnetic bottom was "shrunk" to 2 cm after entering one door and "grown" to 8 cm after entering another door of the machine. In the warm-up phase, two 25 × 32 × 26 cm black boxes featuring a frog and a muffin were used, along with three plastic frogs and three plastic muffins. In the test phase, there was a black "music box" with silver outlines of circle, square, triangle and heart shapes on each of its faces and a remote-controlled doorbell. There were also small (3-cm-wide), medium (6-cm-wide), and large (12-cm-wide) silver circles, squares, and triangles presented as one of each kind, as well as one 3-cm-wide small heart and one 12-cm-wide large heart for the test trials.

2.3. Procedure

In Experiment 1 (see Fig. 1), participants were introduced to a monkey who "likes to wear hats that fit *just right* on his head"; to demonstrate, the experimenter placed the well-fitting blue crown and the green beanie on the monkey's head. The participants then learned about two "Change Machines". The experimenter inserted one red ball and one blue sponge into the machines: one machine produced *bigger* versions of the objects while the other produced *smaller* versions. After inserting and removing each object, the experimenter commented, "Wow! Look what happened!" Many participants smiled and laughed at the shrinking and enlarging of objects. The objects transformed during training were kept beside their respective machines to reduce memory demands.

At test, participants saw a hat that was either too big or too small for the monkey (counterbalanced between participants). The experimenter asked, "Which machine should we put it in to make the hat fit *just right* on his head?" Responses were coded as the first verbal utterance (e.g., "The red one!"), point, or intervention (e.g., picking up the hat and placing it inside one of the machines). If a participant produced more than one conflicting response, the response was coded as the door that the child actually opened. Participants who did not respond to the question after a pause of more than five seconds were asked a second time; the few participants who did not respond to the repeated question were excluded for failure to make a choice. Participants were provided the opportunity to actually solve the problem after they provided their responses.

In Experiment 2 (see Fig. 2), participants sat on their parent's lap facing the frog and muffin boxes for the "warm-up" phase. Experimenter 1 asked the child to put "all the frogs in the frog box and the muffins in the muffin box," preparing them for using the change machines.

Once the warm-up was complete, participants underwent "*causal rule training*." Experimenter 1 introduced the black music box and said, "This is my music box! And you know what? Sometimes, when we put shapes on it, it plays music! Do you want to see how it works?" The experimenter turned the box such that the side with an outline of a (medium-sized) circle faced upward. They produced the three silver circles (small, medium, and large) and said, "Look! Let's see if the circles make it play music." In separate instances, Experimenter 1 placed the smallest and largest circle on top of the box, within the silver outline, paused to listen, and said, "Huh. No music! Hmm. Let's try this other circle!" Then Experimenter 1 placed the medium sized circle that fitted the outline exactly, and Experimenter 2 surreptitiously activated the doorbell in the music box. Experimenter 1 exclaimed, "Wow, music! We made my music box play music! Do you want to try?" Experimenter 1 removed the circle and handed it to the child, who was encouraged to place it on the outline, thus activating the music box again. This procedure was repeated with the "square" and "triangle" sides of the box; in each case,

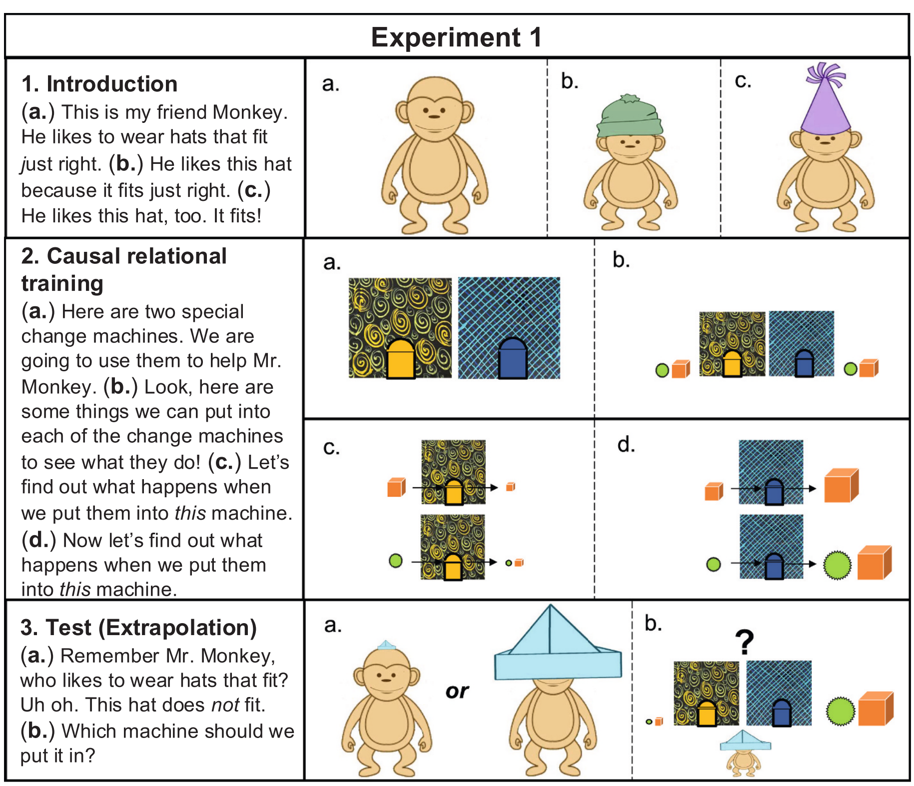

Fig. 1. Stimuli and procedures in Experiment 1. Children were introduced to Monkey, received causal relational training on the change machines, and were tested on their capability to generalize the correct causal relation to a novel hat that can be worn by Monkey.

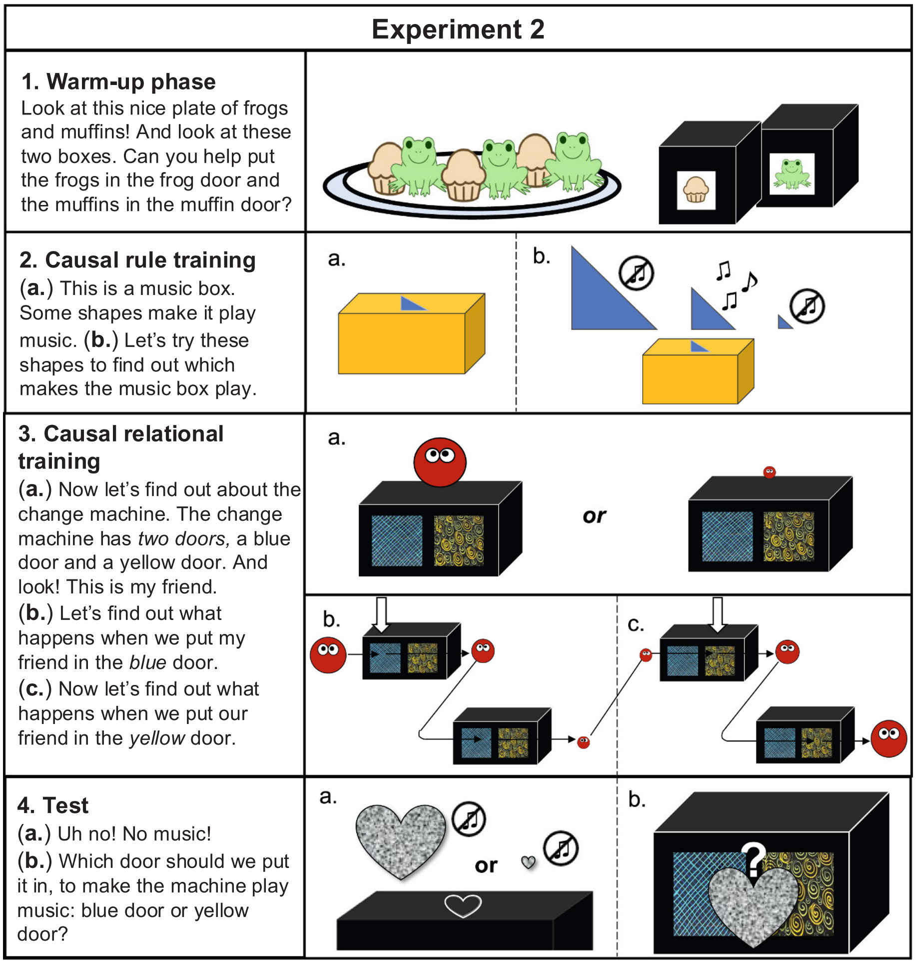

Fig. 2. Stimuli and procedures in Experiment 2. Children underwent a warm-up phase to match frogs and muffins, received a causal rule training involving a music box and a causal relational training on the change machine, and were tested on their capability to generalize the correct causal relation to a novel heart shape so that the music box played music.

only the medium-sized shape fitted the outline exactly and activated the box.

Next, participants underwent "*change machine training*" wherein the training object was placed in each change machine door twice and transformed serially (see Fig. 2, panel 3b and 3c). Experimenter 1 said, "Cool! So that's how my music box works. Now let me show you my other cool machine. This is my change machine. Look! My change machine has two doors. It has a yellow door [*opening and closing the yellow door*] and a blue door [*opening and closing the blue door*]. And look! I have this friend right here [*producing the red spherical character, either in small or large form, counterbalanced*]. Sometimes, when we put my friend inside of the doors, something happens! Let's find out what happens when we put my friend in the [*yellow/blue, counterbalanced*] door." The experimenter opened one of the doors, handed the "friend" to the participant, pointed to the platform inside, and said, "Can you put my friend inside of the [*yellow/blue*] door?" Once the child placed the object on the platform, Experimenter 1 closed the door. Experimenter 2 then spun the rotating platform behind the machine, such that the *big* character was replaced by an identical, but *medium*-sized version, and surreptitiously dinged a bell to signal that the change had occurred. Experimenter 1 said, "Wow! Did you hear that? Let's find out what happened." The door was opened, and the participant was prompted to remove the character. Experimenter 1 said, "Cool! Now, let's put my friend back inside the [*yellow/blue*] *door*." The procedure was repeated, this time with the character emerging in a *small* size.

Once the child had removed the large character from the door, Experimenter 1 said, "Cool! So that's what happens when we put my friend in the [*yellow/blue*] *door*. Now let's find out what happens when we put my friend in the [*blue/yellow*] *door*." The procedure was repeated with the second door—i.e., the participant saw two size changes of the character in the opposite direction. Afterwards, Experimenter 1 said, "Cool! So that's what happens when we put my friend in the [*blue/yellow*] *door*. Ok, let's put my friend away for now!" The order of door presentation and the correspondence between door color and size change were counterbalanced. Again, participants were surprised to see that the objects changed in size.

In the test phase, the experimenter produced the music box and said, "Hey, do you remember my music box? Well guess what. I have *one more shape*—look, it's a heart [*producing either the too-big or too-small heart shape*]! Let's see if we can make my music box play music." The experimenter placed the shape on the heart outline, and the music did not play. The experimenter said, "Hmm. No music," and then asked the child, "Which door should we put it in to make my door play music? Should we put it in the yellow door [*opening and closing the yellow door*] or the blue door [*opening and closing the blue door*]?" After the child provided their answer, the experimenter handed the heart shape to the child and allowed them to open the door of their choice. Responses were coded as in Experiment 1.

3. Results

3.1. Experiment 1

In Experiment 1, three-year-olds ($n = 36$, $M_{age} = 41.3$ months, $SD_{age} = 4.37$ months) observed evidence concerning two "Change Machines." Each participant saw four demonstrations (two for each machine). One 4-cm-wide object at a time was placed out of sight inside of the machine, and a sound effect played. The object was then removed to reveal that it had changed size. The two objects inserted into each machine were perceptually distinct, but the pair of objects for each machine was identical (see Fig. 1). One machine halved the objects' size to 2-cm-wide; the other caused the size to double to 8-cm-wide (see Materials and Methods).

At test, the experimenter reintroduced a character who "likes to wear hats that fit *just right*" (the experimenter never used language referring to size during the experiment). Then, the experimenter presented participants with a novel scenario involving a new object of a new size: 10-cm-wide hat that was either dramatically too large ($n = 18$ participants) or a 2.5-cm-wide hat that was too small ($n = 18$ participants), and asked, "Which machine should we put it in, to make a hat that fits?" Critically, the test object had a different absolute size than the demonstration objects (smaller in the too-small condition, and larger in the too-large condition). In this new scenario, participants had to determine the solution size of the hat in relation to its current novel size. They then had to identify which causal relation they had observed during training that involved very different objects and sizes, to decide how to intervene to produce the novel outcome, which they had never seen. Notably, "perceptual matching" — in this case, size-matching the test hat to the transformed training objects—would yield the *wrong* answer: if participants matched the oversized hat to the machine with enlarged training objects, they would get an even larger hat; conversely, if they matched the undersized hat to the machine that shrank training objects, they would receive an even smaller hat. A significant majority (72.2 %) of participants chose the appropriate machine to solve the problem ($SD = 0.45$, 95 % CI [56.9 %, 87.6 %]), significantly above chance (50 %), $t(35) = 2.94$, $p = 0.006$, $d = 0.49$, with no difference between performance in the "too big" and "too small" conditions, $t(34) = 0.73$, $p = 0.47$, $d = 0.24$. Thus, after only two sets of observations, three-year-olds learned, generalized, and selected the correct causal intervention.

3.2. Experiment 2

Experiment 2 replicated the results of Experiment 1 with a preregistered sample of $n = 100$ toddlers aged 24 to 30 months ($M_{age} = 27.3$ months, $SD_{age} = 2.69$ months). The previous experimental procedure was modified in three ways to accommodate this younger population. First, a "warm-up" phase allowed toddlers to practice picking up and inserting objects behind doors. Second, to reduce the working memory load for toddlers, instead of two machines, there was only one machine with two doors, and instead of multiple objects, a single object was transformed two times *sequentially* for each door to demonstrate the continuous causal function of the doors (see Fig. 2, panel 3b and 3c). For example, half of participants (counterbalanced) saw that the object initially measured 8 cm wide. After passing through the 'shrinking door' once, it reduced to 4 cm, and after two passes, it shrank to 2 cm. Subsequently, passing through the "enlarging door" once increased its width to 4 cm, and after two passes, it expanded back to 8 cm. In contrast to Experiment 1, the transformed object was removed after the demonstrations.

Third, the problem-solving framing ("make a hat that fits") was replaced with a less verbal causal rule learning task (adapted from Sim & Xu, 2017). In an initial demonstration, children observed that shapes that matched the size of silhouette outlines on a box—but not larger or smaller shapes—caused music to play. The "change machine" demonstration followed. At test, children were again presented with the music box, along with a novel heart shape that was either too large (12-cm-wide) or too small (3-cm-wide) to activate it. As in Experiment 1, this test object was both visually dissimilar and different in size from the training objects (see Materials and Methods). The experimenter asked, "Which door should we put it in, to make the machine play music?"

A significant majority of toddlers ($M = 68$ %, $SD = 0.47$, 95 % CI [58.7 %, 77.3 %]) chose the appropriate machine to solve the problem—significantly above chance, $t(99) = 3.84$, $p < 0.001$, $d = 0.38$, with no difference between conditions, $t(98) = 0$, $p = 1$, $d = 0$. When 10 participants who were excluded due to parents' spontaneous use of size language were included in the analysis, performance did not change, $M = 68$ %, $SD = 0.47$, 95 % CI [59.3 %, 77.0 %]; $t(206) = 0.028$, $p = 0.98$, $d = 0.0039$). A logistic mixed effects model with age in months as a dependent variable and a random intercept for each participant showed that age was not a significant predictor of success, $\beta = 0.079$, $SE = 0.084$, $z = 0.95$, $p = 0.34$. Two-year-olds learned, generalized, and actively selected the relevant causal relation to solve a novel problem after only two sets of observations; see Fig. 2.

4. Discussion

The present study demonstrates that "causal relational reasoning" – the ability to learn, generalize, and apply abstract *relations* in novel *causal* interventions – is present very early in life. In two experiments, children as young as 24 months observed a handful of novel causal transformations without any information regarding the mechanism by which the causal event occurred (precluding learning by "intuitive physics"). They were then prompted to solve a novel problem involving a completely unrelated object. Results suggest that children rapidly learned and generalized the abstract relations from a small number of examples, and then readily applied them to choose the appropriate causal intervention that would solve the problem.

These findings provide new evidence for sophisticated abilities for both causal and relational reasoning in toddlers and young preschoolers. Given the mechanistic opacity of the experimental paradigm, participants could not solve the problem by using superficial perceptual generalizations or basic causal reasoning (i.e., merely learning that the machines "made a difference" as opposed to encoding abstract information about the *kind* of differences that they made; Willatts, 2013). Instead, the findings suggest that young children are capable of learning these abstract relations in way that enables them to generalize to different kinds of objects based on just a few observations. Critically, these relations extend beyond the 'same' and 'different' relations in prior studies that have provided evidence for early relational reasoning (Dewar & Xu, 2010; Ferry et al., 2015; Goddu et al., 2020; Hochmann et al., 2016; Walker & Gopnik, 2014).

Moreover, young children appear to have possibly extrapolated their knowledge beyond their observations to identify the causal relation that would yield an appropriate new result. In the training phase of both experiments, children observed that machines caused objects to change size. Crucially, however, in the test phase of both experiments, the test object was a different *absolute* size than any of the training objects previously observed: they were either bigger than the biggest object previously observed, or smaller than the smallest one. Thus, children's success was due to learning the causal relational rules governing the object transformations ("bigger" and "smaller") rather than simple perceptual matching (e.g., categorizing objects as "big" or "small" based on their absolute dimensions). In Experiment 1, three-year-olds were also given the possibility of incorrectly matching "size" with "size." Small objects observed during training were placed next to the shrinking door, while large ones were placed next to the enlarging door. If children were simply matching absolute size of new objects at test with these training objects, then they would have mistakenly chosen to place the too-small hat into the machine that made objects even smaller and the too-large hat into the machine that made objects even bigger. Yet our results suggest that they were not tempted by this possibility. An alternate, slightly more complex form of possible perceptual matching would be one in which the size of the objects that the machine produced were seen as a proxy for the process it performs (e.g., "I need to make the hat *bigger*; I'll put it in the machine with the *big* objects beside it."). However, this strategy was ruled out by Experiment 2, in which even younger children (who saw no objects beside the machines at test) were still able to learn and appropriately apply the size relations.

The fact that children chose the appropriate intervention between two possible machines that transformed objects along the *same dimension*, but in *different directions*, also suggests a deeper grasp of the comparative aspect of size. Future studies might explore the possible nuances of children's causal relational inferences, such as whether what they are learning is something like a continuous mathematical function of size that enables them to extrapolate more generally and precisely. For instance, children might be presented with two machines that both increase objects' sizes, but to varying degrees (e.g., linear versus exponential), or that systematically change featural properties (e.g., the quantity of spots or stripes) on an object in a similar way. Another intriguing research direction might test for the possibility of magnitude matching across domains. For example, would children select the size-increasing box as the one to make a lamp glow brighter, or a musical toy to play louder?

This study shows that toddlers have an impressively early understanding of abstract causal relations, but it also speaks to other questions about early conceptual abilities. In contrast to some earlier claims, these experiments also show that toddlers and preschoolers possess an impressive ability to understand rules (Bunge & Wallis, 2008; Zelazo, 2007) in a sophisticated manner, identifying the relevant and specific dimension of object change that governs the rule and applying the rule to new situation. These results also suggest that two- and three-year-olds can already make some kinds of scale judgments, contrary to early findings that suggested that young children make scale errors (DeLoache, Uttal, & Rosengren, 2004). In these experiments, toddlers could still infer abstract size functions and determine the appropriate function to perform a novel intervention that they had not yet seen. At least in this particular context, very young children seem to understand both rules and scale.

One crucial aspect of these experiments is that the mechanism by which the machines transformed the objects was causally *opaque* to participants. Humans seem to be unique among species in developing tools so advanced that they frequently obscure physical causality (e.g., remote controls and touch screens). This corroborates other research suggesting that children can comprehend abstract principles either concurrently with or even before they learn the specific causal relations that underpin them (Adibpour & Hochmann, 2023; Gopnik & Wellman, 2012; Hernik & Csibra, 2009; Horner & Whiten, 2005; Lucas, Bridgers, Griffiths and Gopnik, 2014; McGuigan, Whiten, Flynn, & Horner, 2007; Rozenblit & Keil, 2002).

More generally, the present study demonstrates that early human causal relational reasoning tracks not only simple causal learning from statistical contingencies to reproduce an already observed outcome (as in the case of "blicket detectors"), but also abstract relations for generating unseen outcomes and solving new problems. This may suggest novel approaches in causal reasoning research, including insights for causal inference in artificial systems. Most work in causal inference has focused on learning and generalizing point-like, binary or probabilistic difference-making relations between variables. In contrast, systems that can learn to attend to the abstract forms of relevant causal relations may better approximate the general form of causal judgments that humans appear to make naturally, from a very young age, in their goal-directed action and problem-solving.

CRediT authorship contribution statement

Mariel K. Goddu: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Eunice Yiu:** Writing – review & editing, Visualization, Validation, Formal analysis, Conceptualization. **Alison Gopnik:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Investigation, Funding acquisition, Conceptualization.

Data availability

We have shared all data from the experiments in Appendix A.

Acknowledgements

We gratefully acknowledge our funding sources: DARPA 047498-002 Co-PI Machine Common Sense, John Templeton Foundation 61475 The Development of Curiosity, Templeton World Charity Foundation PI TWCF0434 Play: A Computational Account, DOD ONR MURI Co-PI Self-Learning Perception through Real World Interaction, Bezos Family Foundation Fund and the Lisa M. Capps Memorial Endowment Award, which funded this research. We also thank the Berkeley School, Monteverde, Step One School, the Bay Area Discovery Museum, the Children's Creativity Museum, and the Lawrence Hall of Science for allowing data collection at their fine facilities. Finally, we thank research assistants Robert Alba, Zoe Appel, Jessica Chung, Emily Demsetz, Zennia Dillon, Erin Fernwood, Madi Hurst, Jack Nelson, Madeleine Levac, Yvette Sanchez, Jocelyn Woo and Charlene Zhu for their creativity and hard work for making the "change machines" come to life.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2024.105959.

References

Adibpour, P., & Hochmann, J. R. (2023). Infants' understanding of the causal power of agents and tools. *Proceedings of the National Academy of Sciences, 120*(50), Article e2309669120.
Adolph, K. E., & Hoch, J. E. (2019). Motor development: Embodied, embedded, enculturated, and enabling. *Annual Review of Psychology, 70*, 141–164.
Ahearne, C., Dilworth, S., Rollings, R., Livingstone, V., & Murray, D. (2016). Touch-screen technology usage in toddlers. *Archives of Disease in Childhood, 101*(2), 181–183.
Buchanan, D. W., & Sobel, D. M. (2011). Mechanism-based causal reasoning in young children. *Child Development, 82*(6), 2053–2066.
Bullock, M. (1984). Preschool children's understanding of causal connections. *British Journal of Developmental Psychology, 2*(2), 139–148.
Bunge, S. A., & Wallis, J. D. (Eds.). (2008). *Neuroscience of rule-guided behavior*. Oxford University Press.
Carey, S., & Spelke, E. (1996). Science and core knowledge. *Philosophy of Science, 63*(4), 515–533.
Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development, 11*(3), 356–373.
Christie, S., & Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science, 38*, 383–397.
Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148–153.
DeLoache, J. S., Uttal, D. H., & Rosengren, K. S. (2004). Scale errors offer evidence for a perception-action dissociation early in life. *Science, 304*(5673), 1027–1029.
Dewar, K. M., & Xu, F. (2010). Induction, overhypothesis, and the origin of abstract knowledge: Evidence from 9-month-old infants. *Psychological Science, 21*(12), 1871–1877.
Ferry, A. L., Hespos, S. J., & Gentner, D. (2015). Prelinguistic relational concepts: Investigating analogical processing in infants. *Child Development, 86*(5), 1386–1405.
Gärdenfors, P., & Lombard, M. (2020). Technology led to more abstract causal reasoning. *Biology and Philosophy, 35*(4), 40.
Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*(2), 155–170.
Gentner, D. (2003). Why we're so smart. In *Language in mind: Advances in the study of language and thought* (pp. 195–235).
Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure mapping and relational language support children's learning of relational categories. *Child Development, 82*(4), 1173–1188.
Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14*(4), 487–513.

Gilhooly, K. J., & Murphy, P. (2005). Differentiating insight from non-insight problems. *Thinking & Reasoning, 11*(3), 279–302.
Goddu, M. K., & Gopnik, A. (2020). Learning what to change: Young children use "difference-making" to identify causally relevant variables. *Developmental Psychology, 56*(2), 275.
Goddu, M. K., & Gopnik, A. (2024). The development of human causal learning and reasoning. *Nature Reviews Psychology*, 1–21.
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.
Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological review, 111*(1), 3.
Gopnik, A., & Sobel, D. M. (2000). Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child development, 71*(5), 1205–1222.
Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental psychology, 37*(5), 620.
Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin, 138*(6), 1085.
Goswami, U. (2013). *Analogical reasoning in children*. Psychology Press.
Hammer, R., Diesendruck, G., Weinshall, D., & Hochstein, S. (2009). The development of category learning strategies: What makes the difference? *Cognition, 112*(1), 105–119.
Hernik, M., & Csibra, G. (2009). Functional understanding facilitates learning about tools in human children. *Current Opinion in Neurobiology, 19*(1), 34–38.
Hochmann, J. R., Mody, S., & Carey, S. (2016). Infants' representations of same and different in match-and non-match-to-sample. *Cognitive Psychology, 86*, 87–111.
Holyoak, K. J., & Thagard, P. (1994). *Mental leaps: Analogy in creative thought*. The MIT Press.
Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition, 8*, 164–181.
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.
Magid, R. W., Sheskin, M., & Schulz, L. E. (2015). Imagination and the generation of new ideas. *Cognitive Development, 34*, 99–110.
McGuigan, N., Whiten, A., Flynn, E., & Horner, V. (2007). Imitation of causally opaque versus causally transparent tool use by 3-and 5-year-old children. *Cognitive Development, 22*(3), 353–364.
Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31*(5), 838.
Paulus, M., Hunnius, S., Van Elk, M., & Bekkering, H. (2012). How learning to shake a rattle affects 8-month-old infants' perception of the rattle's sound: Electrophysiological evidence for action-effect binding in infancy. *Developmental Cognitive Neuroscience, 2*(1), 90–96.
Pearl, J. (2000). *Models, reasoning and inference*. Cambridge, UK: Cambridge University Press.
Pearl, J. (2009). Causal inference in statistics: An overview. *Statist. Surv., 3*, 94–146.
Piaget, J., & Cook, M. (1952). *The origins of intelligence in children*. Vol. 8, No. 5 (pp. 18–1952). New York: International Universities Press.
Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology, 94*(3), 249–273.
Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science, 26*(5), 521–562.
Sim, Z. L., & Xu, F. (2017). Learning higher-order generalizations through free play: Evidence from 2-and 3-year-old children. *Developmental Psychology, 53*(4), 642.
Sloan, A. T., Jones, N. A., & Kelso, J. S. (2023). Meaning from movement and stillness: Signatures of coordination dynamics reveal infant agency. *Proceedings of the National Academy of Sciences, 120*(39), Article e2306732120.
Sobel, D. M., & Legare, C. H. (2014). Causal learning in children. *Wiley Interdisciplinary Reviews: Cognitive Science, 5*(4), 413–427.
Spelke, E. S. (2014). Physical knowledge in infancy: Reflections on Piaget's theory. In *The epigenesis of mind* (pp. 133–169). Psychology Press.
Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science, 10*(1), 89–96.
Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Lawrence Erlbaum.
Tomasello, M. (2009). *The cultural origins of human cognition*. Harvard University Press.
Ullman, T. D., Spelke, E., Battaglia, P., & Tenenbaum, J. B. (2017). Mind games: Game engines as an architecture for intuitive physics. *Trends in Cognitive Sciences, 21*(9), 649–665.
Uzgiris, I. C., & Hunt, J. (1975). *Assessment in infancy: Ordinal scales of psychological development*.
Walker, C. M., & Gopnik, A. (2014). Toddlers infer higher-order relational principles in causal learning. *Psychological Science, 25*(1), 161–169.
Want, S. C., & Harris, P. L. (2002). How do children ape? Applying concepts from the study of non-human primates to the developmental study of 'imitation' in children. *Developmental Science, 5*(1), 1–14.
Willatts, P. (2013). Development of problem-solving strategies in infancy. In *Children's strategies* (pp. 23–66). Psychology Press.
Woodward, J. (2005). *Making things happen: A theory of causal explanation*. Oxford University Press.
Zelazo, P. D. (2007). The development of rule use in childhood. In *Neuroscience of rule-guided behavior*. Oxford University Press.