Empowerment Gain and Causal Model Construction: Children and adults are sensitive to controllability and variability in their causal interventions

Philosophical Transactions of the Royal Society A (In press),

Eunice Yiu, Dept. of Psychology, UC Berkeley
Alison Gopnik Dept. of Psychology, UC Berkeley

Abstract

Learning about the causal structure of the world is a fundamental problem for human cognition. Causal models and especially causal learning have proved to be difficult for Large Models using standard techniques of deep learning. In contrast, cognitive scientists have applied advances in our formal understanding of causation in computer science, particularly within the Causal Bayes Net formalism, to understand human causal learning. In the very different tradition of reinforcement learning, researchers have described an intrinsic reward signal called "empowerment" which maximizes mutual information between actions and their outcomes. "Empowerment" may be an important bridge between classical Bayesian causal learning and reinforcement learning and may help to characterize causal learning in humans and enable it in machines. If an agent learns an accurate causal world model they will necessarily increase their empowerment, and increasing empowerment will lead to a more accurate causal world model. Empowerment may also explain distinctive empirical features of children's causal learning, as well as providing a more tractable computational account of how that learning is possible. In an empirical study, we systematically test how children and adults use cues to empowerment to infer causal relations, design effective causal interventions and appropriately generalize to new contexts.

Learning about the causal structure of the world is a fundamental problem for human cognition, and causal world models are central to intuitive theories. Such models allow wide generalization from limited data. Current large language models (LLMs), despite impressive progress in other areas, have very limited abilities to learn about causation or to perform genuine causal inference (Bender et al., 2021; Jin et al., 2024; Lewis & Mitchell, 2024). The inferential abilities they do display depend on the fact that they detect patterns in text and pictures that are generated by causal model-building humans (Yiu, Kosoy & Gopnik, 2024; Yiu et al., 2024). Such a model might predict that the words "fire" and "smoke" are associated without any conception that fire causes smoke. In contrast, even very young children can and do spontaneously construct novel causal models of the world around them (e.g., Cook et al., 2011; Gopnik et al., 1999, 2004; Schulz & Bonawitz, 2007).

Where do these models come from in humans and how is this kind of learning possible? Over the last twenty years or so, cognitive scientists have applied advances in our formal understanding of causation in philosophy and computer science to understand human causal learning. In particular, researchers have used the Causal Bayes Net formalism which relates directed acyclic causal graphs to patterns of conditional probability, interventions and counterfactuals in systematic ways (Pearl, 2000; Spirtes, Glymour & Scheines, 2000). It provides a natural way to describe both causal models and the patterns of data they generate.

The Bayes net formalism assumes an “interventionist” account of causation – roughly, variable X is causally related to variable Y if an intervention changing the value of X would lead to a change in the value of Y. This has also become a dominant account of causality in philosophy. Causation is distinctive, on this view, because it allows for successful action. Simply observing correlations or associations between events may allow you to predict one event from another. But an intervention to produce one event will only lead to another if there is a causal relation between them. For example, yellow nicotine stained fingers, as well as smoking, may be correlated with lung cancer, and you might predict that people with yellow fingers are more likely to develop cancer. But intervening to wash off hands won’t affect the cancer while intervening to stop smoking will. The Bayes net formalism allows you to appropriately predict the consequences of interventions on variables as well as predicting the association between them.

This allows the reverse Bayesian inference, at least in principle – from the pattern of interventions and associations it should be possible to infer the causal model that was most likely to have generated that data. This approach to causal learning is part of a broader Bayesian approach to learning in developmental cognitive science (for a recent review see Ullman & Tenenbaum, 2020). Empirically, causal learning in children is remarkably well captured by this formalism (for reviews see Gopnik & Wellman, 2012; Gopnik & Bonawitz, 2014; Goddu & Gopnik, 2024). Given a pattern of data, even preschool children infer the correct causal hypothesis.

From a computational perspective, however, inferring causal models from evidence, like many other kinds of Bayesian inference, proves to be intractable – the space of possible hypotheses is too large. Various computational techniques have been used to try to deal with this problem, particularly sampling methods, and there is some evidence that young children use similar methods (Bonawitz et al., 2014a, 2014b). but the search problems remain challenging. Another approach might be to suggest that potential causal models are strongly constrained by prior innate “core knowledge” (e.g., Spelke, 2022). Although this might help address the search problem, it also undermines one of the major advantages of causal learning. Causal learning is so valuable precisely because it allows us to go beyond innate knowledge and learn about causal structure that may be counter intuitive – such as the workings of a TV remote or a smart phone, or eventually of scientific physics or psychology.

Another challenge for both Bayesian approaches and the classic “deep learning” algorithms that drive large language models involves the role of active exploration and experimentation in causal learning. Children are not simply passive consumers of data. Instead from very early on they actively seek out evidence that is relevant to the causal problems they are trying to solve, and this exploration plays an important role in their solutions to those problems. There has recently been extensive and elegant work showing just how motivated children are to experiment and explore, and how intelligently they do so (e.g., Schulz, 2012; Giron et al., 2023). Similarly, in formal science, experimentation is the canonical way to discover causal structure. But, with a few exceptions, (e.g., Eberhardt & Spirtes, 2007) there have not been computational accounts of how this kind of active intervention and experimentation takes place and how it allows causal inference

However, in parallel, an apparently quite different kind of learning mechanism – reinforcement learning – has become increasingly influential in neuroscience and computer science, particularly when combined with modern “deep” machine learning (Sutton & Barto, 2018; Dayan & Balleine, 2002). For example, deep reinforcement learning has been the key to DeepMind's accomplishments in mastering Go and Chess (Silver et al., 2018).

As the philosopher James Woodward, among others, has pointed out, classical reinforcement learning might be thought of as a very specific and narrow form of causal learning (Woodward, 2007, 2023). In particular, it relates the specific actions an agent performs on the world – their interventions – to specific outcomes, in the form of external rewards. This basic structure is similar to the basic structure of causal relations on the interventionist account underlying the causal Bayes net formalism. Moreover, RL contrasts with other types of learning that are more purely associative, such as classical conditioning or neural network learning mechanisms. Associative learning captures correlations between variables and allows predictions about those variables, but doesn't allow a special role for interventions, and so is further removed from causal learning.

However, reinforcement learning is much narrower in application than causal learning more generally, it typically applies only to the agent's own actions and to the rewards that follow those actions. More importantly, the basic motivational structure of reinforcement learning and causal learning are very different. Reinforcement learning is motivated by utilities – the attempt to maximize external rewards. Classically, causal learning is epistemically motivated – it involves approximating the true structure of the world. It is true that ultimately causal knowledge allows effective action and decision-making, and so increases utilities. But this is a long term and indirect effect. In the Bayesian framework, decision-making and utility calculations are layered on top of the fundamental epistemic project (as in "causal decision theory"). The reverse is true in reinforcement learning – RL agents may try to learn about the environment but they only do so in service of the fundamental utility project – that is, they learn in order to maximize further rewards later on.

Although RL methods can be very effective in relatively low-dimensional and fixed environments such as go or chess, especially with dense reward signals, they have much more difficulty in the high-dimensional and open-ended environments that are characteristic of human causal learning. In particular, RL systems face consistent difficulties in balancing exploitation -- the fundamental utilitarian motivation of maximizing reward, and exploration -- the more epistemic motivation of learning about the structure of the environment (e.g., Sutton & Barto, 2018; Cohen et al., 2007). In the long run, exploration will lead to more effective action and reward, but it requires the agent to forego reward in the short run. These problems arise even in very simple "bandit tasks" where agents are only required to choose between actions with known or unknown outcomes. Choosing the known option allows an agent to be sure of reward, but the unknown option will be more informative and lead to more knowledge and so to more effective action in the long run.

One approach to solving these problems has been to propose systems that seek internal epistemic rewards in addition to the classical external rewards. These systems implement the intrinsic epistemic motivation of classical causal learning in an RL framework. Several types of intrinsic rewards have been proposed in the literature. They include a variety of curiosity-based rewards, particularly measures of information gain, entropy and novelty (Oudeyer et al., 2007; Schmidhuber, 2010; Pathak et al., 2017). These rewards do seem to increase the exploratory efficacy of reinforcement learning systems. Moreover, there is evidence that children and even infants also seek out such intrinsic rewards, particularly information gain (Schulz, 2012; Kidd & Hayden, 2015).

However, it is still difficult to sufficiently constrain these systems. For example, a classic failure mode for systems that seek out information gain is the "noisy TV" problem (Schmidhuber, 2010). Such systems will be captured by random noise like TV static. Seeking novelty and information gain by themselves doesn't seem to be adequate to understand the environment effectively in a way that supports action.

A different approach uses "empowerment" as an intrinsic epistemic reward (Klyubin et al., 2005, 2008). Interestingly "empowerment" originally was formulated in the evolutionary computation and artificial life literature as part of an attempt to specify how intelligence might emerge even in very simple organisms. The crucial idea is that intelligence involves systematic relationships between "sensors" -systems for sensing and perceiving the environment, and "actuators" -systems for acting on the environment. A number of evolutionary theorists have pointed out that brains emerged in concert with the emergence of animals with coordinated perceptual and motor systems – such as eyes and claws -- in the Cambrian explosion. This contrasts, with complex, large and successful organisms in the earlier Ediacaran period that do not have these features, and do not have brains. (Godfrey-Smith, 2020: Jablonka & Lamb 2014)). Empowerment described how sensors and actuators could be coordinated in an adaptive way.

In empowerment, an agent maximizes the mutual information between its actions and their outcomes, regardless of the reward value of those outcomes. In other words, the system is rewarded if variation in an action systematically leads to parallel variation in the outcome so that the value of the action predicts the value of the outcome. Simultaneously, the system is also rewarded for maximizing the variety of actions it takes, ensuring that it explores the widest range of possible actions and outcomes. Thus seeking empowerment leads to both actions that allow more control of the environment and more variable actions.

Recently this idea has been applied to a variety of reinforcement learning problems (Du et al., 2020; Zhao et al., 2019; De Abril & Kanai, 2018). By endowing RL agents with empowerment as an intrinsic reward those agents can explore and represent the environment more effectively. Rather than simply seeking out novelty or information, such a system will seek out exactly the relations in the world that involve the closest match between actions and outcomes – the most controllable relations. This means that it will discover the relations that will be most likely to be useful for a wide range of future goal-directed actions. If I discover, for example, that moving a stick systematically changes the position of objects that it contacts, I can later pick up the stick to draw an out of reach object towards me. In fact, infants seem to learn to use sticks in just this way (Uzgiris & Hunt, 1975).

Although they come from different traditions, we argue that causal learning and empowerment gain are intimately related. In particular, if an agent learns an accurate causal model of the world they will necessarily increase their empowerment, and, vice versa, increasing empowerment will lead to a more accurate (if implicit) causal world model.

This claim is rooted in the peculiar nature of causation. In the past, we and others have argued that it is helpful to think of causal learning as an "inverse problem" (Gopnik et al., 2004). An inverse problem involves inferring the structure of the external world from the data that world generates. A classic example is the way a visual system infers the structure of the three-dimensional world from retinal images or pixels. Other examples might include the way that we infer beliefs and desires from behavior in "theory of mind", infer neural structure from fMRI evidence in cognitive neuroscience, or infer atomic structure from cloud chamber traces. In all these cases we assume that there is, in fact, some single distinctive structure in the outside world that we are trying to reconstruct. A God's eye view of the universe would discover something corresponding to the 3-d structure or psychological structure or neural structure or atomic structure that we are trying to reconstruct. This structure is independent of the agents who are trying to understand it. Bayesian approaches in cognitive science essentially formalize these inverse problems in a probabilistic way.

However, the interventionist accounts that underlie causal Bayes nets (Woodward, 2007; Pearl, 2000; Spirtes, Glymour & Scheines, 2000) imply a different relation between agents and the world. The asymmetries between cause and effect that are central to interventionist accounts of causation are, notoriously, hard to be find in physics, at least at the micro level. Instead, our notions of causation are rooted in the idea that causal relations are precisely those external relations that support an agent's interventions. Philosophers and computationalists in this tradition define causal relations as those relations such that intervening to alter the value of a cause variable will lead to a corresponding change in the effect variable – in short, those relations where actions predictably produce outcomes. The ideal interventions that underpin causal inference are not identical to the intentional actions of actual agents but they are closely related, and in many circumstances agents' intentional actions will also serve as ideal interventions (see Woodward, 2020), and so pick out causal relationships in the world. Discovering such relations between interventions and outcomes is also the fundamental idea behind empowerment. But there may not be any single God's eye view agent-independent causal structure analogous to, say, 3-d spatial structure or atomic structure.

The philosopher Peter Godfrey-Smith (2009), among others has argued for "causal pluralism" – where many ontologically disparate phenomena can all support causal interventions. On this view there is nothing analogous to the spatial structure of the 3-d world in the causal case. Rather there are many quite different relationships in the world, ranging from the intuitive physical relations that support "billiard-ball" causation to the belief-desire relationships of intuitive psychology, to the highly counter-intuitive relations of physics, that all happen to systematically support causal interventions. Again these are precisely the relations of control that would produce the greatest increases in empowerment. Maximizing empowerment will lead to the discovery of causal relations, and vice-versa.

A further distinctive feature of empowerment, is that it can have both "mind to world" and "world to mind" directions of fit. It is possible to maximize an agent's empowerment by increasing its knowledge about how the world works (matching the mind to the world as in science). This is the classic picture of causal learning. However, it is equally possible to maximize empowerment by increasing an agent's skill and control over the world (matching the world to the mind as in engineering). This is more like the orientation of classic reinforcement learning.

However, whether we think in terms of science or engineering, the basic structure of causal relations will be the same, intervene systematically on X to systematically change the value of Y. This is well captured by the notion of empowerment. Of course, in more abstract and conceptual cases, the interventions may be theoretical rather than actual – to say that the moon causes the tides is to say that if we could alter the position of the moon the tides would also alter, even if this intervention isn't actually possible. But, if a relation is genuinely causal this sort of intervention should at least be conceptually possible. Moreover, this distinguishes causal relationships from other relationships such as spatial, geometric or logical relationships. And, in practice, the test for whether we have discovered causal relations is to perform experiments –to determine whether experimentally varying one variable will systematically predict the value of another, that is, precisely to look for high mutual information between interventions and outcomes.

Thinking about empowerment might also help us understand the psychology of causal learning. A recent paper suggests that adult's exploration of a video-game like environment can be well characterized by empowerment (Brändle et al., 2023) and we have shown that this is also true for children (Du et al., 2023). But the empowerment approach more generally captures important features of early causal knowledge and learning and helps to explain a wide range of developmental findings.

Looking-time studies suggest that very young infants perceive some particular relations in intuitive physics that support causal inference, such as the relations of movement and collision in "billiard-ball" causality (Leslie, 1982). However, the development of causal concepts more broadly is initially closely linked to actual goal-directed actions on the world and their outcomes. We have suggested (Goddu & Gopnik, 2024, following Woodward, 2000, 2020) that both in phylogeny and ontology, causal understanding moves from a first-person perspective, to a third-person perspective to an impersonal perspective. Reinforcement learning, which is found in almost all intelligent animals, represents a causal relation between the animal's own actions and their outcomes. Imitation learning, which is found in some non-human animals in limited ways, but is ubiquitous in human infants, represents a causal relation between another animals' actions and outcomes. The sort of impersonal causal understanding in science represents causal relations in the world independent of actual actions, though crucially supporting such actions in principal. Empowerment may be applied to all three types of relations.

Is there empirical evidence that humans including young children seek something like empowerment, and that this contributes to their causal learning? In the 70's, interestingly in the context of thinking about operant conditioning and reinforcement learning, a series of papers suggested that even very young infants are indeed rewarded by something like empowerment. In classic studies of "conjugate reinforcement" Rovee-Collier (1979) tied a ribbon from a crib mobile to the infant's foot, so that kicking made the mobile move. Infants as young as 3 months old systematically acted to make the mobile move, varying their actions and observing the correlation between those actions and the behavior of the mobile. There were similar results in studies where infants could make a mobile move or activate a pattern of lights by turning their heads on a pressure sensitive pillow (Watson, 1972; Papousek & Papousek, 1975). Moreover, these actions could not simply be explained by classic reinforcement learning with the mobile's motion as a reward. Infants would learn to turn their heads or kick and would continue to act to do so, even though they no longer looked at the mobile or the lights, aside from a brief glance to check that their action was effective. Infants varied their actions and observed their results rather than simply converging on a single effective action. In addition, infants smiled and cooed when their actions consistently led to an effect, but not when that effect simply occurred independently of their actions (Watson, 1972). A more recent study with this methodology shows that infants consistently alter their actions on the mobile in a way that increases the contingency of their actions, again unlike classical reinforcement learning (Sloan et al., 2023). In fact, Rovee-Collier described her results precisely as an empowerment reward: "*The control which the infants have gained over the consequences of their own actions seems to have become the reward, rather than the specific consequences per se.*" (Rovee-Collier, 1979).

In "conjugate reinforcement" infants are acting to maximize the empowerment of their own actions – their causal understanding has a first-person perspective. We know that from early in infancy children also represent the goal-directed actions of others and distinguish them from other kinds of events and movements (e.g., Woodward, 1998, 2009). Moreover, they map the goal-directed actions of others on to their own actions (Meltzoff, 2007). This is an important feature of human causal learning. It distinguishes it from other types of learning, such as classic reinforcement learning, which only concern an agent's own actions, and also distinguish it from learning in other animals (Taylor et al., 2014). From early in life, then, children have the cognitive and conceptual structures in place to discover empowerment relations between actions and outcomes, both their own and those of others.

From at least 24 months and probably earlier, children make genuine causal inferences by observing the correlations between goal-directed actions – their own or others—and outcomes. However, until around age 4, they do not make similar inferences from simple correlations between events (Wasimeyer & Meltzoff, 2017; Bonawitz et al., 2010; Meltzoff, Waismeyer & Gopnik, 2012). Suppose a 24- month-old sees a human hand repeatedly push a toy car against a block A, which causes a light to go on. Pushing the car against another block B does not have this effect. Now we ask the infant to make the light appear. Infants will reproduce the correct action on A in order to make the light go, but not the action on B. However, they will not do this if they simply see the car move on its own and cause the effect. This is true even though they will look towards the light in this condition, suggesting that they have learned the correlation between the motion of the car and the light (Meltzoff, Waismeyer & Gopnik, 2012). In short, toddlers appear to detect empowerment relations between actions and outcomes and use those relations to infer causal relationships that determine their own future interventions. They do not do this based on correlations among events that do not involve actions and outcomes. 4-year-olds do infer new interventions from correlations alone, but this ability seems to depend on their earlier learning through goal-directed action.

Empowerment also naturally applies to children's early exploratory play (Chu & Schulz, 2020). Even infants characteristically play by varying their actions on an object and observing the results – hence the perennial popularity of toys like rattles and busy boxes that afford such empowering actions. Empowerment based reinforcement learning, unlike Bayesian inference, also provides a natural way to characterize such experimental actions, they are precisely what you would expect from a system that was trying to act to maximize empowerment.

Thinking of causal learning in terms of empowerment may also help to resolve some of the search problems. Maximizing empowerment would not require the sort of search through a high-dimensional hypothesis space that is so challenging for Bayesian inference. Unfortunately, precisely calculating mutual information itself poses problems of tractability – but some very recent approximation methods make such calculations more feasible (e.g., Zhao et al., 2020). Children might also begin by simply looking for correlations between actions, their own and others, and the outcomes that follow them, rather than fully calculating mutual information.

If children are maximizing empowerment they would have a mechanism for independently discovering causal relations that are not specified innately, even without requiring the full apparatus of Bayesian causal inference. They might look for relations that have the feature of mutual information between their own actions and those of others and outcomes, like the relations between sticks and distant toys. This might then allow them to build up a repertoire of basic causal arrows that can then be combined to build more complex models.

Testing Empowerment Empirically – The Star Machines

We described a number of developmental studies and observations suggesting that children may indeed be seeking empowerment gain in their everyday play and exploration. But could we test this idea more systematically? Empowerment involves two factors. Actions must lead to effects in the environment – they must enable control, but simply seeking control might lead you to mindlessly repeat the same action with a deterministic outcome over and over. To obtain high empowerment, actions must also be variable, and this variability should be correlated with variability in the environment. Again this contrasts with simply seeking novelty or information in the environment. Just as the failure case for simple control would be a deterministic loop, the failure case for simple novelty seeking would be to pursue randomly generated novel outcomes, as in the "noisy TV" case we described earlier. To maximize empowerment, and to explore the causal structure of the environment most effectively, you should seek a combination of control and variability.

In the following experiments, we systematically contrast empowerment gain, which involves both controllability and variability, with novelty, that is simple variability, on the one hand, or efficacy, that is simple control, on the other. A rotary dimmer dial is more empowering than a two-way light switch because it allows fine-tuned control over many more distinct brightness levels, rather than just two fixed states. However, if that same rotary dial functioned like a wheel of fortune, randomly determining an arbitrary brightness level with each turn, it would be less empowering than the basic light switch because despite providing a variety of outcomes, it lacks predictable control. Similarly, the rotary dial would yield more causal knowledge than the switch or the wheel of fortune. Will children and adults differentiate these cases and make appropriate causal inferences, interventions and generalizations as a result? Will they prefer events that afford empowerment over similar events with lower controllability or variability? We also ask whether children and adults have different preferences and make different interventions in some contexts and not others.

Study Method

To assess if humans appreciate this distinction, we designed a study to examine whether children recognize and prefer control and / or variability when they make causal generalizations and interventions. We introduced 80 five- to ten-year-old children (μ = 7.52 years, *SD* = 1.68 years) and 120 adults (μ = 27.57 years, *SD* = 4.30 years) to three machines, each designed to generate outcomes reflecting variability, controllability, or a combination of both. In a cover story participants were told that "the elf boss wants you to make stars with these machines". The machines were set up so that placing an object, initially a star, in one of several slots produced another different object (see Figure 1).

(a) Demonstration Phase

We first demonstrated how each machine produced outcomes. One machine (the purely controllable machine) always produced the same star size, regardless of which slot the demonstrator used. It either generated a large star or a medium star (60 adults and 46 children were randomly assigned to observe big stars from this machine, while the rest were assigned to the medium-star condition). Another machine (the controllable and variable machine) had a perfect correlation between slot size and output size: the big slot produced a big star, the medium slot produced a medium-sized star (unchanged from the input size), and the small slot produced a small star. The third machine (the purely variable machine) generated star sizes randomly, with no correlation to the slot size. The machines were otherwise identical, and the color and position of the machines were randomized across participants.

Participants were not told about the underlying causal structure of the machines, but had to infer them by observing the narrator drag stars into different slots on each machine. Every time a star was generated by the machine, a narrator commented on the change in size compared to the input in one of the following ways: "Look it becomes smaller!", "Look it becomes bigger!", "Look it is the same!". All participants watched three outputs per slot, amounting to twenty-seven outcomes observed across the three three-slot machines. Figure 1 presents an example of what the participants saw in the demonstration. To eliminate participants' need to recall the outcome patterns generated by the three machines, the star outputs remained visible on the screen throughout the experiment.

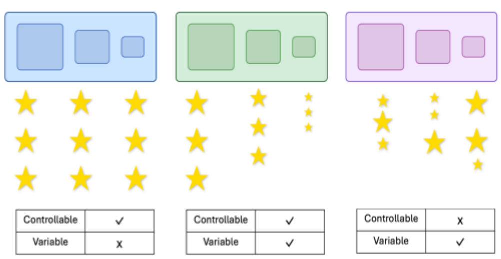


Figure 1. Three machines characterized by the controllability and variability of their outputs. (left) The purely controllable machine generates a single, deterministic output across all slots. (middle) The controllable and variable machine produces three distinct outputs, each reliably corresponding to slot size. (right) The purely variable machine generates three different outputs in a completely stochastic manner, with no predictable pattern. The color and order of the machines were fully randomized across participants.

(b) Generalization Tests

After observing the demonstration, participants were asked to design interventions to solve new causal problems. These interventions required generalization at various levels of abstraction. The first level of abstraction involved generalizing the structure of the machine to a new output value. Participants were introduced to an "extra small" slot that was newly appended to each of the three machines and were asked to generate an "extra small" star that would be smaller than those they had previously observed. The correct choice was the extra small slot in the variable and controllable machine (as opposed to the extra small slots in the other two machines; chance level = 1/3), as it was the only one that would reliably produce the new outcome.

The second level of abstraction required generalization to a new object. Participants were given a new object, hats, and were told to make nine different-sized hats: three large, three medium and three small using any slot from any of the three machines (twelve slots in total). To make big hats, one could put hats in any of the four slots in the controllable machine if it exclusively produces large outcomes, and / or the large slot in the controllable and reliable machine (chance level = 5/12 if the former is true; chance level = 1/12 if the latter is true). To make medium-sized hats, one could pass hats into any of the four slots in the controllable machine if it exclusively produces medium outcomes, and / or the medium slot in the controllable and reliable machine (chance level = 5/12 if the former is true; chance level = 1/12 if the latter is true). To make small hats, one could place hats into only the small or extra small slots in the controllable and reliable machine (chance level = 2/12). So the combined chance level of making different sized hats is 25% for participants in the condition where the controllable machine makes only big stars as well as in the condition where the controllable machine makes only medium-sized stars.

The third level of abstraction demanded generalization from object size to a new perceptual dimension – brightness. Participants were given light bulbs and were told that the light bulbs could be made bright, sort of bright, sort of dim and dim with the machines. They were then asked to make a bright light bulb and a dim light bulb. This requires the same solution as making a big hat (chance level = 5/12 or 1/12) and a small hat (chance level = 2/12) in the previous problem.

(c) Machine Preference

The experiment concluded by asking participants which of the three machines they would keep if they were asked to "*work* to make new things" or if they could simply be "given more things to *play* with."

Study Results

(a) Generalization

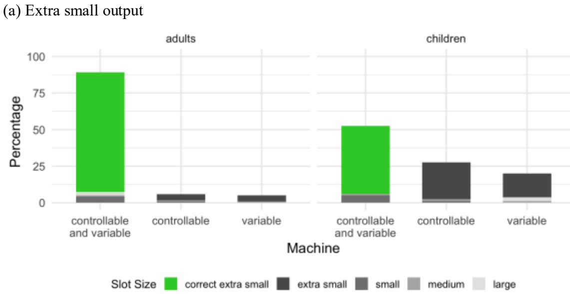
When asked to generalize to a new output value by making an extra small star, 46.25% of children and 80.83% of adults correctly selected the extra small slot on the variable and controllable machine (see Figure 2(a)). This is significantly above chance (1/3) for both children ($p = .017$) and adults ($p < .001$) on a binomial test. Chi-square tests show the variable and controllable machine was significantly preferred by both children ($\chi2(2) = 13.9$, $p < .001$) and adults ($\chi2(2) = 166.4$, $p < .001$). There was no significant difference between the other two machines.

When asked to generalize to a new object by producing hats of different sizes, children selected the correct slots on the controllable machines 56.26% the time (*SE* = 3.35%) and adults 86.85% the time (*SE* = 1.90%) (see Figure 2(b)). Both children ($t(79) = 9.03$, $p < .001$) and adults ($t(119) = 32.62$, $p < .001$) performed significantly above chance level (25%). Across the nine trials, both groups showed a significant preference for the variable and controllable machine over the purely variable machine (children: $z = 6.24$, $p < .001$; adults: $z = 15.02$, $p < .001$) and over the purely controllable machine (children: $z = 3.15$, $p = .001$; adults: $z = 9.30$, $p < .001$). The preference for pure variability over pure controllability was significant only in children ($z = -3.09$, $p = .002$) but not in adults ($z = -5.72$, $p < .001$).

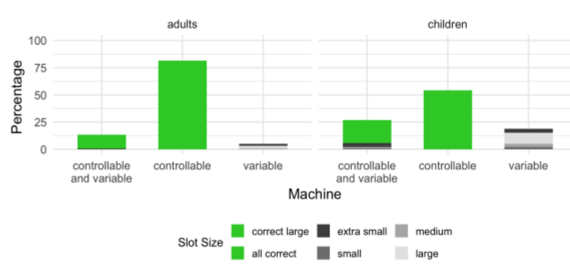
When asked to generalize to a new perceptual modality by creating light bulbs that alter in brightness instead of size, children and adults again succeeded in selecting the appropriate slots in the controllable machines. In the condition where the purely controllable machine also made big objects only, 71.74% children and 76.67% adults successfully created bright light bulbs, significantly above the chance level of 5/12 (binomial test, $p < .001$) (see Figure 2(c)(i)); in the condition where the purely controllable machine made medium sized objects, 58.33% children and 70% adults made bright light bulbs, also again significantly above the chance level of 1/12 (binomial test, $p < .001$) (see Figure 2(c)(ii)). Both children and adults (aggregated over the conditions where the purely controllable machine generates only medium outcomes or only large outcomes) showed significant preference for the variable and controllable machine over the other two machines on a Chi-squared test (children: $\chi 2(2) = 10.2$, $p = 0.0061$; adults: $\chi 2(2) = 54.95$, $p < .001$) (see Figure 2(c)(iii)). When making dim light bulbs, 48.10% children and 68.91% adults selected the correct slots, which again is significantly above the chance level of 2/12 (binomial test, $p < .001$). As in making bright light bulbs, both children and adults significantly preferred the variable and controllable machine on a Chi-squared test (children: $\chi 2(2) = 17.4$, $p < .001$; adults: $\chi 2(2) = 108.12$, $p < .001$). Across the trials, there was no significant difference in usage between the purely variable machine and the purely controllable machine.

Even though the purely variable machine could generate the full range of outcomes demanded by the three generalization tasks, children and adults rarely used this machine because these outcomes are unpredictable and they have no control over them (20% and 5.04% respectively in making an extra small star; $\mu = 22.64\%$, $SE = 2.22\%$ and $\mu = 7.78\%$, $SE = 1.24\%$ in making nine different sized hats; $\mu = 22.15\%$, $SE = 3.31\%$ and $\mu = 15.06\%$, SE = 2.32% in making bright and dim light bulbs). This highlights humans' appreciation for controllability over variability – more specifically, the systematic variation between inputs and outputs – in generalization and designing interventions for novel outcomes. It is interesting, however, that children showed more of a preference for pure variability than adults. This may simply reflect the greater noise in the children's responses, but it may also suggest that children are especially sensitive to information gain.

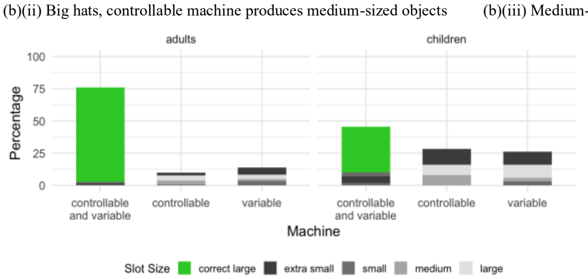
(a) Extra small output



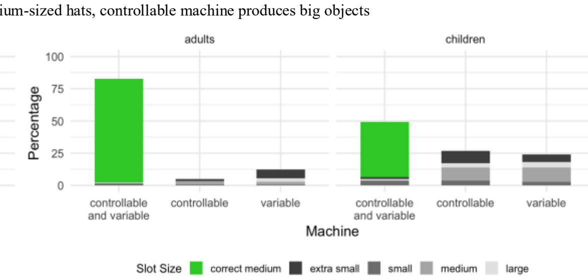
(b)(i) Big hats, controllable machine produces big objects



(b)(ii) Big hats, controllable machine produces medium-sized objects



(b)(iii) Medium-sized hats, controllable machine produces big objects



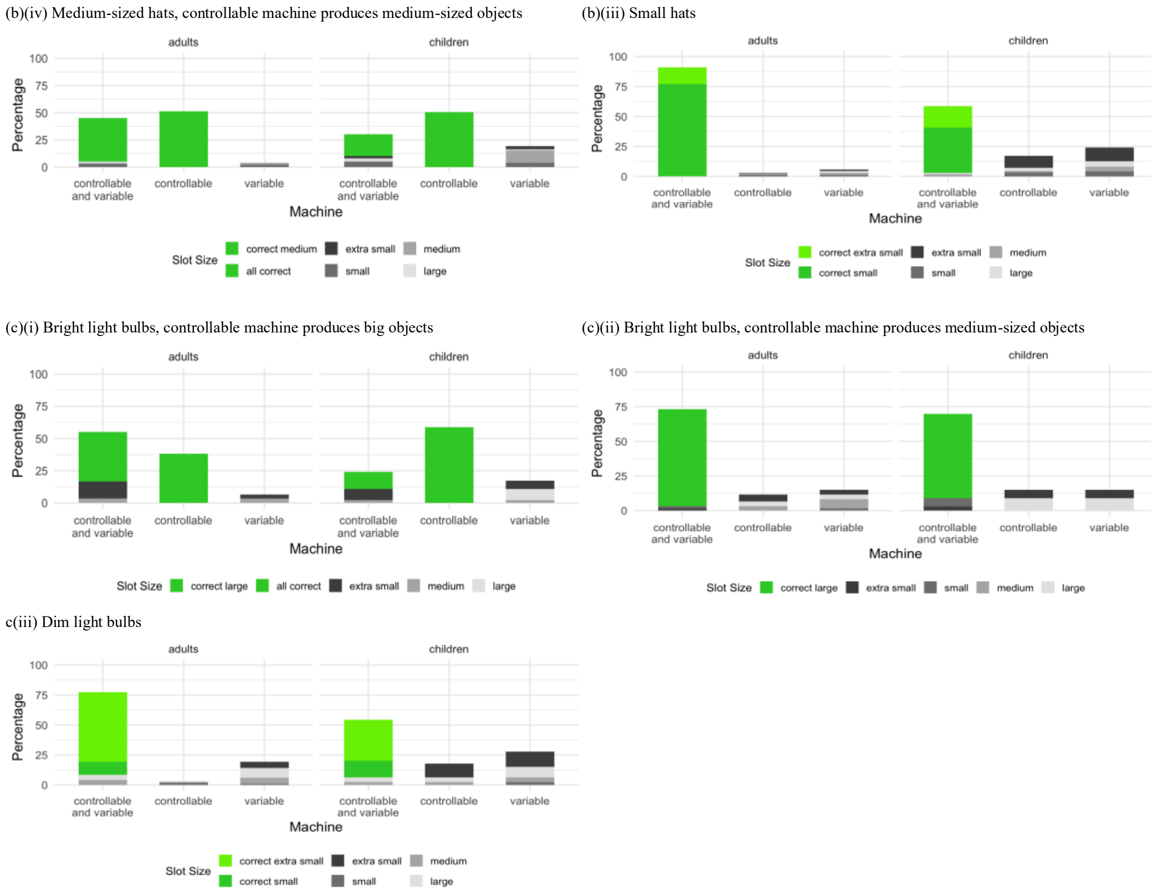


Figure 2. Distribution of machine and slot preference across the three generalization tasks: (a) to a new output value (extra small star), (b) to a new object kind (hats), (c) to a new perceptual modality (brightness).

(b) Preference

When asked to select a machine to keep for work to make new kinds of things, the variable and controllable machine is most preferred (selected by 48.75% children and 75% adults), while the purely variable is least preferred (selected by 20% children and 66.7% adults) (see Figure 3, left). A Chi-squared test revealed that children favored the variable and controllable machine (z = 2.93), preferred the purely controllable machine less (z = -0.40, not significant), and avoided the purely variable machine (z = -2.53), $p = .009$. Adults showed an even stronger preference for the controllable and variable machine (z = 9.68) and the other two machines were significantly under-selected (z = -3.49 for the purely controllable machine and -6.20 for the purely variable machine, $p < .001$ for both).

When asked to select a machine to keep for playing more, adults but not children continued to exhibit a strong preference for the controllable and variable machine (selected by 59.17% adults and 31.25% children) (see Figure 3, right). Children's choices were evenly distributed on a Chi-squared test (variable and controllable: z = -.39, variable: z = .32, controllable: z = .079). By contrast, adults still strongly preferred the variable and controllable machine (z = 6.00) to the purely variable machine (z = -3.10, p < .001) and the purely controllable machine (z = -2.90, $p < .001$).

There was a significant shift in adults' machine preferences between the work and play contexts, $\chi2(2) = 10.94$, $p = .012$. Adults were significantly more likely to select the variable and controllable machine in the work context compared to the play context, $\chi2(1) = 5.68$, $p = .017$, while their preference for the purely variable machine increased in the play context, $\chi2(1) = 8.65$, $p = .0033$. There was no significant change in the selection of the controllable machine. Likewise, children demonstrated a significant shift in preference from controllability in the work context to variability in the play context, $\chi2(3) = 6.04$, $p = .014$. Like adults, children were significantly more likely to select the variable and controllable machine in the work context compared to the play context, $\chi2(1) = 6.04$, $p = .014$. At the same time their preference for the purely variable machine strengthened in the play context, $\chi2(1) = 5.04$, $p = .025$. There is no difference in preference for the purely controllable machine between large and medium-sized outcomes in both children and adults, indicating that object size does not explain their preference. In summary, children and adults show a stronger preference for the controllable and variable machine in the work context and shift to prefer the purely variable machine more in the play context.

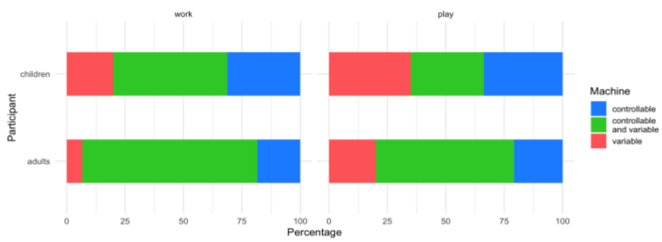


Figure 3. Proportion of machine selections by children and adults in the work and play contexts.

Study Discussion

In a simple setup featuring otherwise identical systems with different levels of control and variability, we find that both children and adults can successfully leverage controllability and variability to design interventions that generate novel, unseen causal outcomes. Whether producing a star of a completely new size, creating new hats with different sizes, or adjusting a new perceptual dimension – the brightness of light bulbs – participants reliably selected the correct slots on the controllable and variable machine (and sometimes the controllable but not variable machine when the target outcome did not involve variation in value).

To achieve various specific goals at different levels of generalization, children and adults needed to use a machine that embodied empowerment. They needed to not just observe the direct mapping between a single input and a single output, but also determine if variation in a slot systematically leads to parallel variation in the outcome, so that the action of putting an object into a slot predicts the value of the outcome. As we described above, this is the fundamental characteristic of causal relationships (Gopnik, 2024). Notably, participants rarely relied on the variable but uncontrollable machine, even though this machine could generate the same diversity of outcomes as the variable and controllable machine.

Explicit preferences for machine use in work versus play conditions further illustrate children's and adults' sensitivity to controllability. Children and adults showed a strong preference for the controllable and variable machine in goal-directed work, with adults showing a stronger preference. In contrast, the purely variable machine was more often chosen for play, possibly due to a desire to resolve uncertainty or find unpredictability appealing when there is no goal. Nevertheless, empowerment represented by controllable variability in this study, emerged as a key tool for goal-directed causal generalization and intervention.

Conclusion

In sum, we argue here that recent work on "empowerment" may help bridge Bayesian and RL approaches to learning and provide both empirical and theoretical insight into the crucial problem of learning the causal structure of the world. We also present a first set of experiments exploring this idea empirically.

Acknowledgements and Support: This research was supported by grants from the DARPA Machine Common Sense Program, the ONR MURI program, the Templeton Foundation, the Templeton World Charity Foundation and CIFAR. Yu Qing Du, Li Dayan and Eliza Kosoy were central to these ideas and the empirical research and discussions with Pieter Abbeel, Lerrel Pinto, Yann LeCun, Tom Everitt and the members of BAIR and the Ranch Metaphysics Workshop all contributed greatly to clarifying these ideas.

References

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).

Bonawitz, E. B., Ferranti, D., Saxe, R., Gopnik, A., Meltzoff, A. N., Woodward, J., & Schulz, L. E. (2010). Just do it? Investigating the gap between prediction and action in toddlers' causal inferences. *Cognition*, *115*(1), 104-117.

Bonawitz, E., Denison, S., Griffiths, T. L., & Gopnik, A. (2014). Probabilistic models, learning algorithms, and response variability: sampling in cognitive development. *Trends in cognitive sciences*, *18*(10), 497-500.

Bonawitz, E., Denison, S., Gopnik, A., & Griffiths, T. L. (2014). Win-Stay, Lose-Sample: A simple sequential algorithm for approximating Bayesian inference. *Cognitive psychology*, *74*, 35-65.

Brändle, F., Stocks, L. J., Tenenbaum, J. B., Gershman, S. J., & Schulz, E. (2023). Empowerment contributes to exploration behaviour in a creative video game. *Nature Human Behaviour*, *7*(9), 1481-1489.

Chu, J., & Schulz, L. E. (2020). Play, curiosity, and cognition. *Annual Review of Developmental Psychology*, *2*, 317-343.

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 933-942.

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

Dayan, P., & Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron*, *36*(2), 285-298.

De Abril, I. M., & Kanai, R. (2018). A unified strategy for implementing curiosity and empowerment driven reinforcement learning. *arXiv preprint arXiv:1806.06505*.

Du, Y., Tiomkin, S., Kiciman, E., Polani, D., Abbeel, P., & Dragan, A. (2020). Ave: Assistance via empowerment. *Advances in Neural Information Processing Systems*, *33*, 4560-4571.

Du, Y., Kosoy, E., Dayan, A., Rufova, M., Abbeel, P., & Gopnik, A. (2023, November). What can AI Learn from Human Exploration? Intrinsically-Motivated Humans and Agents in Open-World Exploration. In *Neurips 2023 workshop: Information-theoretic principles in cognitive systems*.

Eberhardt, F., & Scheines, R. (2007). Interventions and causal inference. *Philosophy of science*, *74*(5), 981-995.

Giron, A. P., Ciranka, S., Schulz, E., van den Bos, W., Ruggeri, A., Meder, B., & Wu, C. M. (2023). Developmental changes in exploration resemble stochastic optimization. *Nature Human Behaviour*, *7*(11), 1955-1967.

Godfrey-Smith, P. (2009). Causal pluralism. In H. Beebee, C. Hitchcock, & P. Menzies (Eds.), *The Oxford Handbook of Causation*. Oxford University Press.

Godfrey-Smith, P. (2020). *Metazoa: Animal life and the birth of the mind.* Farrar, Straus and Giroux.

Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (1999). *The scientist in the crib: Minds, brains, and how children learn*. William Morrow & Co.

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

Gopnik, A. (2024). Empowerment as Causal Learning, Causal Learning as Empowerment: A bridge between Bayesian causal hypothesis testing and reinforcement learning.

Goddu, M. & Gopnik, A. (2024). The development of human causal learning and reasoning. *Nature Reviews Psychology*.

Jablonka, E., & Lamb, M. J. (2014). *Evolution in four dimensions, revised edition: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. MIT press.

Jin, Z., Chen, Y., Leeb, F., Gresele, L., Kamal, O., Lyu, Z., ... & Schölkopf, B. (2024). Cladder: A benchmark to assess causal reasoning capabilities of language models. *Advances in Neural Information Processing Systems*, *36*.

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron*, *88*(3), 449-460.

Klyubin, A. S., Polani, D., and Nehaniv, C. L. (2005). Empowerment: A universal agent-centric measure of control. In Evolutionary Computation, 2005. The 2005 IEEE Congress on (IEEE), vol. 1, 128–135.

Klyubin, A. S., Polani, D., & Nehaniv, C. L. (2008). Keep your options open: An information-based driving principle for sensorimotor systems. PloS one, 3(12), e4018.

Leslie, A. M. (1982). The perception of causality in infants. *Perception*, *11*(2), 173-186.

Lewis, M., & Mitchell, M. (2024). Using counterfactual tasks to evaluate the generality of analogical reasoning in large language models. *arXiv preprint arXiv:2402.08955*.

Meltzoff, A. N. (2007). 'Like me': a foundation for social cognition. *Developmental science*, *10*(1), 126-134.

Meltzoff, A. N., Waismeyer, A., & Gopnik, A. (2012). Learning about causes from people: observational causal learning in 24-month-old infants. *Developmental psychology*, *48*(5), 1215.

Oudeyer, P. Y., Kaplan, F., & Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE transactions on evolutionary computation*, *11*(2), 265-286.

Papoušek, H., & Papoušek, M. (1975, January). Cognitive aspects of preverbal social interaction between human infants and adults. In *Ciba Foundation Symposium 33-Parent-Infant Interaction* (pp. 241-269). Chichester, UK: John Wiley & Sons, Ltd..

Pathak, D., Agrawal, P., Efros, A. A., & Darrell, T. (2017, July). Curiosity-driven exploration by self-supervised prediction. In *International conference on machine learning* (pp. 2778-2787). PMLR.

Pearl, J., & Mackenzie, D. (2018). *The book of why: the new science of cause and effect*. Basic books.

Pearl, J. (2000). *Causality*. Cambridge university press.

Piaget, J., & Cook, M. (1952). *The origins of intelligence in children* (Vol. 8, No. 5, pp. 18-1952). New York: International Universities Press.

Rovee-Collier, C. K., & Gekoski, M. J. (1979). The economics of infancy: A review of conjugate reinforcement. *Advances in child development and behavior*, *13*, 195-255.

Salge, C., Glackin, C., & Polani, D. (2014). Empowerment–an introduction. *Guided Self-Organization: Inception*, 67-114.

Schulz, E., Wu, C. M., Ruggeri, A., & Meder, B. (2019). Searching for rewards like a child means less generalization and more directed exploration. *Psychological science*, *30*(11), 1561-1572.

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

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

Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE transactions on autonomous mental development*, *2*(3), 230-247.

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., ... & Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, *362*(6419), 1140-1144.

Spelke, E. S. (2022). *What babies know: Core knowledge and composition volume 1* (Vol. 1). Oxford University Press.

Spirtes, P., Glymour, C. N., & Scheines, R. (2000). *Causation, prediction, and search*. MIT press.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Taylor, A. H., Cheke, L. G., Waismeyer, A., Meltzoff, A. N., Miller, R., Gopnik, A., ... & Gray, R. D. (2014). Of babies and birds: complex tool behaviours are not sufficient for the evolution of the ability to create a novel causal intervention. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1787), 20140837.

Ullman, T. D., & Tenenbaum, J. B. (2020). Bayesian models of conceptual development: Learning as building models of the world. *Annual Review of Developmental Psychology*, *2*, 533-558.

Uzgiris, I. C., & Hunt, J. (1975). Assessment in infancy: Ordinal scales of psychological development.

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

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1-34.

Woodward, A. L. (2009). Infants' grasp of others' intentions. *Current directions in psychological science*, *18*(1), 53-57.

Woodward, J. (2007). Interventionist theories of causation in psychological perspective. *Causal learning: Psychology, philosophy, and computation*, 19-36.

Woodward, J. (2021). *Causation with a human face: Normative theory and descriptive psychology*. Oxford University Press.

Yiu, E., Kosoy, E., & Gopnik, A. (2024). Transmission versus truth, imitation versus innovation: What children can do that large language and language-and-vision models cannot (yet). *Perspectives on Psychological Science*, *19*(5), 874-883.

Yiu, E., Qraitem, M., Wong, C., Majhi, A. N., Bai, Y., Ginosar, S., ... & Saenko, K. (2024). Kiva: Kid-inspired visual analogies for testing large multimodal models. *arXiv preprint arXiv:2407.17773*.

Zhao, R., Abbeel, P., & Tiomkin, S. (2020). Efficient online estimation of empowerment for reinforcement learning. *arXiv preprint arXiv:2007.07356*.