

Developmental Psychology

Children's Exploration Distinguishes Necessary and Merely Possible Outcomes

Luisa Andreuccioli, Sophie Mazor, Katarina Begus, Elizabeth Bonawitz, Stephanie Denison, and Caren M. Walker

Online First Publication, September 21, 2026. <https://dx.doi.org/10.1037/dev0002269>

CITATION

Andreuccioli, L., Mazor, S., Begus, K., Bonawitz, E., Denison, S., & Walker, C. M. (2026). Children's exploration distinguishes necessary and merely possible outcomes. *Developmental Psychology*. Advance online publication. <https://dx.doi.org/10.1037/dev0002269>

Children's Exploration Distinguishes Necessary and Merely Possible Outcomes

Luisa Andreuccioli¹, Sophie Mazor², Katarina Begus³, Elizabeth Bonawitz⁴,
Stephanie Denison⁵, and Caren M. Walker²

¹ Department of Cognitive Sciences, Central European University

² Department of Psychology, University of California, San Diego

³ Department of Psychology, University of Copenhagen

⁴ Graduate School of Education, Harvard University

⁵ Department of Psychology, University of Waterloo

Preschoolers often fail tasks that require that they select between a necessary and a merely possible outcome. Prior work has interpreted these failures as evidence that children hold a minimal representation of possibility or as reflecting performance factors. However, all previous paradigms have operationalized children's modal understanding as a discrete choice or outcome, making it difficult to distinguish between these accounts. We introduce a continuous measure of search behavior to assess how modal status regulates exploration. American 3- and 4-year-olds (48% female, 52% male) searched for an object dropped into one of two boxes after passing through a set of either transparent (yielding a single necessary location) or opaque tubes (yielding two possible locations). In Experiment 1, both age groups spent relatively less time searching the first location they approached and switched more often when the outcomes were merely possible than when the outcome was necessary. Experiment 2 replicated this pattern in a touchscreen version. Experiment 3 demonstrated the same effect while controlling for object tracking by equating inferential demands across conditions. Across experiments, children's search behavior was aligned with modal status, indicating that representations of necessity and possibility guide exploration before children reliably succeed on discrete choice modal reasoning tasks. These results suggest children's early sensitivity to modal concepts is expressed in their continuous search behavior, shaping real-time decision making. Our findings inform theories of modal reasoning in early childhood, suggesting that modal representations may be more readily expressed through graded patterns of exploration rather than discrete choices.

Public Significance Statement

Young children often struggle with tasks that require explicit reasoning about possibilities. Using a novel paradigm that tracked children's exploration rather than their discrete choices, we found that 3- and 4-year-olds successfully adapted their search behavior for an object depending on whether its location was necessary (one location) or merely possible (two potential locations). These results reveal an early sensitivity to this distinction, offering new insight into the development of children's understanding of multiple possibilities before they can express it in language.

Keywords: cognitive development, modal reasoning, knowledge representations, alternative possibilities

Judith H. Danovitch served as action editor.

Luisa Andreuccioli  <https://orcid.org/0009-0003-1737-2898>

Sophie Mazor  <https://orcid.org/0009-0009-4014-888X>

Katarina Begus  <https://orcid.org/0000-0002-0264-4534>

Elizabeth Bonawitz  <https://orcid.org/0000-0002-7005-9345>

Stephanie Denison  <https://orcid.org/0000-0002-6658-4139>

Caren M. Walker  <https://orcid.org/0000-0001-6532-8611>

Luisa Andreuccioli and Sophie Mazor shared first authorship. Anonymized data will be available upon publication at https://researchbox.org/4711&PEER_REVIEW_passcode=HQGTHO. Ethics approval for this study was granted by the University of California, San Diego's Institutional Review Board. The authors declare no conflicts of interest.

This work was funded by the National Science Foundation Early Career Award (Grant SBE 2047581) awarded to Caren M. Walker. Because this research was funded in whole, or in part, by the Templeton World Charity Foundation (Grant TWCF-2024-33964), for the purpose of open access, the author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

The authors thank the Fleet Science Center, Alexis Burke, Aarthi Popat, and preschools in the greater San Diego area for help with recruitment. The authors are grateful to Brian Leahy, Susan Carey, and members of the Early Learning and Cognition Lab for helpful discussions.

Luisa Andreuccioli played a supporting role in formal analysis and an equal role in conceptualization, investigation, methodology, visualization, and writing—original draft. Sophie Mazor played a lead role in data curation and formal analysis and an equal role in investigation, methodology, visualization, and writing—original draft. Katarina Begus played a supporting role in conceptualization, funding acquisition, and writing—review and editing. Elizabeth Bonawitz played a supporting role in conceptualization, funding acquisition, and writing—review and editing. Stephanie Denison played a supporting role in conceptualization, funding acquisition, and writing—review and editing. Caren M. Walker played a lead role in conceptualization, funding acquisition, project administration, resources, supervision, and writing—review and editing, a supporting role in investigation and methodology, and an equal role in writing—original draft.

Correspondence concerning this article should be addressed to Sophie Mazor, Department of Psychology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, United States. Email: smazor@ucsd.edu

The ability to reason about multiple possibilities is central to human cognition. Not only is it key to everyday decision making (Fodor, 1975; Simon, 1955), but it is also fundamental to the very process of learning: We consider possibilities when evaluating explanations (Lombrozo, 2006), inferring causal relationships (Gopnik & Welman, 2012), thinking counterfactually (Hegarty, 2004; Kahneman & Tversky, 1981), planning for the future (Beck et al., 2006), and making logical inferences from initial premises (Johnson-Laird, 2010; Johnson-Laird et al., 1992). Importantly, reasoning about possibilities is not only a matter of making explicit judgments but also impacts how learners allocate attention, effort, and action when outcomes are uncertain.

On some accounts, the ability to represent alternative possibilities is relatively late-emerging and dependent on the acquisition of modal logic—that is, symbolic operators that allow the learner to distinguish what is necessary from what is merely possible (Leahy & Carey, 2020). In line with this proposal, children under 4 years of age appear to consistently fail tasks that require them to explicitly reason about alternative possibilities (Kim et al., 2016; Leahy, 2023; Leahy et al., 2022; Mody & Carey, 2016; Redshaw & Suddendorf, 2016; Rohwer et al., 2012). For example, in a task in which children are asked to pick between a single cup that is guaranteed to contain a reward and two cups that are merely possible reward locations, 3-year-olds only pick the “safe bet” about half the time. This behavior is also observed in modified versions of this task that maintain the same logic: one in which working memory demands are reduced (Leahy, 2023) and one involving additional scaffolding (Leahy et al., 2022). Similar failures have also been observed in tasks that require children to prepare for mutually exclusive future outcomes (Beck et al., 2006; Leahy, 2024; Redshaw & Suddendorf, 2016; Robinson et al., 2006). For example, when asked to catch an object that is dropped from an inverted Y-shaped tube, young children typically only cover a single exit, rather than using two hands to prepare for both possible outcomes (Redshaw & Suddendorf, 2016). Some have interpreted these findings to mean that young learners are unable to distinguish between necessity and possibility until they have acquired modal language (e.g., “may” vs. “have to”), which typically emerges around 4 years of age (Leahy & Carey, 2020; Leahy & Zalnieriunas, 2021; Mody & Carey, 2016; Ozturk & Papafragou, 2015). Critically, however, these paradigms have all operationalized modal understanding as a discrete choice or outcome, leaving open the possibility that early-emerging modal concepts shape children’s reasoning before they are evident in such responses.

Indeed, a growing body of empirical evidence suggests that the foundations for modal reasoning are in place much earlier (Alderete & Xu, 2023; Baumann et al., 2024; Cesana-Arlotti et al., 2018, 2022, 2025; Goupil et al., 2016; Stahl & Feigenson, 2024; Téglás & Bonatti, 2016; Téglás et al., 2007; Turan-Küçük & Kibbe, 2024, 2025), supported by early-developing systems for probabilistic inference (e.g., Alderete & Xu, 2023), logical quantification (e.g., Cesana-Arlotti et al., 2025), or other mechanisms. Infants as young as 6 to 12 months can integrate probability information to infer the most likely outcome from among multiple alternatives, indicating an early sensitivity to multiple possibilities (Denison & Xu, 2010, 2014, 2019), and by 14 months, infants’ pupillometry responses suggest they are holding multiple possibilities in mind (Cesana-Arlotti et al., 2022). Twelve-month-olds also succeed in deductive reasoning tasks (*a* or *b*; not *a*; therefore *b*) that require them to

represent mutually exclusive object identities (Cesana-Arlotti et al., 2018, 2020), and 18–30-month-olds can infer multiple possible causes from the same information (Goddu et al., 2021). Specifically, when toddlers observed ambiguous evidence that was consistent with two equally likely causal rules (i.e., either a single blue cube or pairs of “different” blocks activate a toy), they inferred and flexibly applied both possibilities, seemingly holding the two options in mind simultaneously. These findings suggest early sensitivity to multiple possibilities; however, some have argued that success on these tasks may not require explicit modal representations (Leahy & Carey, 2020).

One way to resolve this puzzle is to recognize that many existing modal reasoning tasks conflate children’s representation of multiple possibilities with how they allocate exploration under uncertainty. If so, apparent failures in forced-choice tasks may arise not from a lack of modal sensitivity but from how uncertainty-driven exploration is expressed under different task demands. Indeed, a large body of work shows that even infants preferentially attend to events that have the potential to reduce uncertainty (Bergus & Bonawitz, 2024; Bergus et al., 2016; Kidd et al., 2012, 2014; Perez & Feigenson, 2022; Stahl & Feigenson, 2015), and young children actively explore when their knowledge is ambiguous, insufficient, or incompatible with their observations (Bonawitz et al., 2012; Cook et al., 2011; Lapidow & Bonawitz, 2023; Lapidow et al., 2022; Liquin & Lombrozo, 2020; Ruggeri et al., 2017; Schulz & Bonawitz, 2007; Sim & Xu, 2017; Wang et al., 2021). In fact, children will often forgo the opportunity to obtain a tangible reward in order to gain new information about the world (Dahmani et al., 2025; Lapidow & Walker, 2020; Liquin & Gopnik, 2022). For instance, in a bandit-like exploration task, where participants could either exploit and maximize immediate rewards or explore at a cost, children as young as 3 are more likely to explore than older children and adults (Blanco & Sloutsky, 2021; Liquin & Gopnik, 2022; Sumner et al., 2019). Indeed, the intrinsic motivation to explore is considered a fundamental characteristic and driver of human learning (Goddu & Gopnik, 2024; Gopnik, 2000, 2024; Rovee-Collier & Gekoski, 1979; Rule et al., 2024).

It is therefore plausible that tasks that present children with a forced choice between a necessary and a merely possible outcome might similarly promote (costly) exploration behavior, even when it means losing the opportunity for a tangible reward. After all, a “merely possible” outcome is an uncertain outcome, and selecting it provides a unique opportunity for information gain. If so, children’s prior failures may reflect their greater motivation to learn about (rather than avoid) instances when multiple possibilities are present. Crucially, this implies that forced-choice paradigms have limitations for detecting modal representations because they collapse uncertainty-driven exploration into a binary response. We propose a shift away from forced-choice tasks that conflate distinct cognitive processes (e.g., uncertainty monitoring, exploration value, action planning) and instead provide a continuous measure of search time that treats modality not as a choice between options but as a factor that shapes how search effort is distributed over time.

This perspective also reframes theoretical debates about early modal reasoning, including the minimal representation account (Leahy & Carey, 2020). Whereas a modal representer can simultaneously hold multiple possibilities in mind, a minimal representer only considers one “possibility” at a time. According to this account, young children and nonhuman animals lack the symbolic marker

that allows them to conceptualize an event as merely possible. Therefore, when faced with multiple possibilities, minimal representers simulate just one outcome and treat this simulation as if it were a matter of fact, until and unless such simulation proves to be false. For example, early success on tasks involving disjunctive inference or object disambiguation (e.g., [Cesana-Arlotti et al., 2018](#)) can occur if children simulate a single outcome (e.g., guessing which of two objects is inside a cup) and then revise the simulation after receiving additional evidence, resulting in longer looking at inconsistent events. Relatedly, early changes in pupil dilation (e.g., [Cesana-Arlotti et al., 2022](#)) could reflect the cognitive demands associated with revising an initial guess in response to negative evidence, rather than discriminating among multiple possibilities.

Over the past 5 years, the minimal account has been widely debated, with several lines of evidence emerging and challenging its claims ([Alderete & Xu, 2023](#); [Baumann et al., 2024](#); [Cesana-Arlotti et al., 2025](#); [Turan-Küçük & Kibbe, 2024, 2025](#)). For instance, some have proposed that children's failure to prepare for mutually exclusive future outcomes may result from specific task demands—that is, difficulties associated with action planning, rather than an inability to engage in modal reasoning. Indeed, when 3- and 4-year-old children observed two actors performing either an optimal action (i.e., placing both hands beneath the two exits of the Y-shaped tubes) or a suboptimal action (i.e., placing only one hand beneath a single exit), they correctly identified the actor who performed the optimal action as the one who would “catch the marble for sure.” Further, those children who had the opportunity to observe the actors' actions were more likely to subsequently perform the optimal action themselves, despite never observing the outcome of either action in advance ([Turan-Küçük & Kibbe, 2024](#); also see [Phillips & Kratzer, 2024](#); [Turan-Küçük & Kibbe, 2025](#)).

Others have suggested that different cognitive demands associated with standard modal reasoning tasks may similarly mask early competence. For instance, [Alderete and Xu \(2023\)](#) argued that occlusion might hinder performance, and they provide evidence that young children will succeed on a disjunctive reasoning task when they have continuous visual access to possible outcomes. The authors also propose that assessing the likelihood of possibilities across different trajectories and locations—for example, tracking a marble's possible paths as it travels through a branching tube—may introduce additional challenges to simulation. In line with this, a recent study by [Turan-Küçük and Kibbe \(2025\)](#) shows that 3- and 4-year-olds successfully prepare for two possible outcomes when they are required to reason about mutually exclusive object identities (i.e., which objects may emerge), rather than possible locations or trajectories (i.e., where an object may emerge).

[Cesana-Arlotti et al. \(2025\)](#) further challenged the minimal view, arguing that preschoolers can discriminate modal categories (i.e., necessity vs. possibility; possibility vs. impossibility) by applying logical quantification over possible outcomes—treating events as necessary when all alternatives lead to the same results, possible when some do, and impossible when none do. In their paradigm, children helped a character navigate probabilistic devices that released her down several ramps. By selectively blocking ramps and placing either desirable outcomes (cupcakes) or harmful outcomes (peppers) at different exits, the researchers created conditions where outcomes were necessary, possible, or impossible. They found that 3-year-olds reliably preferred the logically stronger option (e.g., a guaranteed cupcake over a merely possible one), suggesting that

children evaluated the full set of alternatives rather than simulating just one outcome. Importantly, children's performance also varied with the motivational context (e.g., avoiding peppers vs. gaining cupcakes), raising the possibility that children's behavior in modal reasoning paradigms may sometimes reflect the presence of competing goals, rather than the absence of modal concepts. Together, these findings suggest that young children's modal reasoning is highly task-dependent.

Against this background, we introduce a method designed to isolate modal representations from exploration and reward. In the present studies, we remove the requirement that children pick either a necessary or possible location, avoiding the potential for conflicting motivations to influence children's performance. We created a novel modal reasoning task, drawing upon prior search paradigms (e.g., [Feigenson & Carey, 2003](#); [Van de Walle et al., 2000](#)), in which we replace this forced choice with a continuous measure of children's exploration behavior to assess whether young children's search strategy differs when responding to distinct modal concepts. Critically, our task isolates modal reasoning while holding reward structure and task success constant, allowing search dynamics to serve as a direct index of children's modal representations. Across three experiments, we present 3- and 4-year-olds with two sets of inverted Y-shaped tubes with both exits attached to opaque search boxes. All children observed two trials in which they watched a target object being dropped into the tubes and were asked to locate the object inside the attached boxes. Children could explore either or both boxes and switch freely between the search locations. In a within-subject design, we manipulated the opacity of the inverted Y-shaped tubes across the two trials: In the merely possible trial (fully opaque tubes), participants could not view the trajectory of the object, leading to two equally possible search locations. In the necessary trial, participants either had direct perceptual access (Experiments 1 and 2; fully transparent tubes) or could infer (Experiment 3; half-opaque tubes) the trajectory of the object, leading to a single necessary location.

The existing accounts of the development of modal reasoning make distinct predictions, not about children's explicit judgments, but about how they should distribute search effort over time. If children are able to represent both search boxes as merely possible locations in the fully opaque trial, they should spend significantly less time searching in the first location they approach in this condition, compared to when they observe the object dropped in the fully transparent and half-opaque trials. If, on the other hand, children hold a minimal representation of possibility, they should only consider one outcome in both necessary and merely possible trials, simulating a single trajectory, and treating that simulation as knowledge. We should therefore find no difference in children's search behavior between the two trial types, since the location of the object is treated as “known” in both cases.

Experiment 1

In Experiment 1, we presented 3- and 4-year-old children with a fully opaque and fully transparent set of inverted Y-shaped tubes, each emptying into two opaque search boxes ([Figure 1](#)). In a within-subjects design, children were asked to locate the target object after it was dropped into one of the sets of tubes across two trials. When the target object was dropped into the transparent tubes, children could observe the trajectory of the ball and direct their search to the

correct box to retrieve the object. When it was dropped into the fully opaque tubes, children had no information about where the object fell, leading to two possible search locations. Importantly, unbeknownst to the participant, the boxes were designed to trap the target object before it landed inside, creating an impossible search task. This allowed us to compare the total amount of time children were willing to spend searching in each location for each set of tubes. This experiment was not preregistered.

Method

Participants

A total of 54 children were tested, including 28 3-year-olds (16 females, $M_{\text{age}} = 3.5$ years; range = 3;0–3;11) and 26 4-year-olds (10 females, $M_{\text{age}} = 4.5$ years; range = 4;0–4;11). Sample sizes across all experiments were based on three previous studies investigating modal reasoning (Leahy, 2023; Leahy et al., 2022; Mody & Carey, 2016), as the present work serves as a conceptual replication of prior work using a novel paradigm. Post hoc sensitivity analyses indicated that we were well-powered (.98) to detect effects at $\alpha = .05$ (critical $t = 2.01$). Children were recruited from the lab's participant database and preschools in the broader San Diego area. Ethics approval for this study was granted by the University of California, San Diego's Institutional Review Board. Children were either monolingual English speakers or bilingual speakers with English as a primary language. An additional 10 children (nine 3-year-olds and one 4-year-old) were tested but excluded (one withdrawal, one failure to attend, one failure on familiarization, three experimenter errors, three stimuli failures). Twenty-six participants reported racial demographic information, identifying as White (61.5%), multiracial (30.7%), and Asian (7.6%), and 22 families reported their average yearly income, with 17 families (77%) earning over \$100,000.

Materials

Materials included a 14 mm yellow-painted wooden bead (the target object), stickers, and two sets of inverted Y-shaped tubes, built from polyvinyl chloride pipes (Figure 1) and attached to an inclined plane. One set of tubes was opaque, and the other set was transparent. The top opening was divided inside to allow the

experimenter to control the side that the target object would fall into when dropped. This divider was not visible to participants.

Six black, rectangular, cardboard search boxes were built for the study. The side of the box facing the child included an opening covered by a felt flap. This allowed the child to reach in with one hand to explore the contents without visual access. On the opposite side of the box (facing the experimenter), a hidden drawer-like opening allowed the experimenter to secretly push a target object inside the search box (during introductory and familiarization phases). The search boxes included a fabric tube extending from the top so they could be attached to the bottom of each of the openings of the inverted Y-shaped tubes during test trials. A layer of fabric inside this extension prevented the target object from landing in the boxes. Each box was filled with 10 multicolored pom-poms, which ensured that children had to manually search inside the box to find the target object, which was distinct in its tactile features from the filler objects.

Finally, a tablet was used as a 31-s timer. The tablet played a musical track that increased in tempo to signal that time was running out.

Procedure

The child sat across the table from the experimenter and gave verbal assent to begin the game. A mitten was placed on the child's nondominant hand to ensure they could only search inside one box at a time. Children were instructed to only use their free hand when prompted to search and were reminded of this as needed.

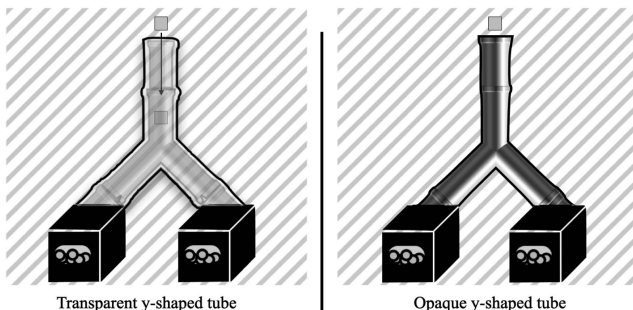
Introductory Phase. The introductory phase was intended to familiarize children with the items involved in the manual search task. The experimenter presented one of the search boxes and told the child that it was filled with soft, fluffy balls (filler objects). Children had the opportunity to touch the fluffy balls when they were outside the box and when they were inside the box. Next, they were asked to hold the yellow wooden ball (target object) and report whether it felt "soft or hard" to draw their attention to the difference between the filler objects and the target object. Children were told that their goal was to "find as many hard balls as possible" and that they would win a sticker every time they found one. The experimenter then dropped the target object from an opening at the top of the box and asked the child to use their dominant hand to find it and remove it.

Familiarization With Search Boxes. The experimenter presented two new search boxes and placed them side-by-side in front of the child. The experimenter said, "There is already a hard ball inside one of these boxes, and I am going to ask you to find it. You can search this box, this box, or both boxes (touching each box in turn)." On the first familiarization trial, the experimenter covertly pushed the target object inside whichever box the child approached first.

Following this, the experimenter brought out two new search boxes, and the same procedure was repeated. This time, however, the experimenter pushed the target object inside whichever box the child did not approach first, so they would have to spontaneously switch to this alternate location to find it. If the child stopped searching before finding the target object, they were given a series of increasingly suggestive prompts. Prompts were only given as needed and presented in the following order: (a) "Have you looked all around the box?" (b) "Remember, you will win a sticker if you can find the hard ball!" (c) "There is definitely a hard ball inside one of these boxes." (d) "You can use all the time you need to find it!"

Figure 1

Illustration of the Two Sets of Inverted Y-Shaped Tubes (With Search Boxes Attached), Presented as Separate Within-Subject Trials in Experiment 1



If children expressed frustration for not being able to find it in a box, and after using all other prompts, we moved onto (e) “Where else could it be?” This final prompt was only needed for three children.

On the next two familiarization trials, the experimenter presented two new search boxes and repeated the same procedure. In these trials, a timer was introduced, and the experimenter instructed the child that they had to find the hard ball before the music stopped to win a sticker. Again, the target object was covertly hidden in the first location the child approached. Then, on the subsequent timed trial, it was hidden in the location that the child did not initially search. After each familiarization trial, the child was given positive or neutral feedback, depending on whether they found the ball before the timer ran out. Children had to find the target object in at least one of the four familiarization trials to be included in the final sample. Only one child failed to do so.

Familiarization With Y-Shaped Tubes. The experimenter introduced the transparent set of Y-shaped tubes, explaining that they would drop the ball from the top of the tubes. They emphasized that children would be able to see the ball as it fell down one of the tubes. The child was asked to point to the tube that the object fell from. The object was dropped four times in a pseudorandom order [L-R-R-L]. Next, the experimenter introduced the opaque set of Y-shaped tubes, emphasizing it would not be possible to see the ball as it fell down the tube. Again, the ball was dropped four times in a pseudorandom order [R-L-L-R], and children were asked to point to the tube that produced the ball.

Test Phase. During the test trials, the experimenter presented either the transparent or opaque set of inverted Y-shaped tubes (order counterbalanced) and said,

Remember those boxes we played with earlier? I am going to attach them to the ends of the tubes. These boxes only have soft balls inside them! Now, I will drop a hard ball from up here, and I am going to ask you to find it.

As in the familiarization trials, the experimenter reminded the child that they had to find the hard ball before the timer ran out to win a sticker.

Critically, the target object never fell into either box, and the child never found the ball on either test trial. Use of the timer ensured a plausible explanation for their inability to locate the ball. Children were given neutral feedback upon failing to find the ball and encouraged to try again. On the second test trial, the same procedure was repeated with the second set of tubes. For both test trials, we recorded time (in seconds) spent searching the first location, time spent in each location, and the total number of times they switched during a given trial. On the transparent trial, we also recorded whether children started searching in the correct (observed) location. One 3-year-old and two 4-year-olds began searching in the incorrect location. These children were included in our primary analyses.¹

Children’s search times were coded using Datavyu (<https://Data.vyu.org>), a software package that allows for frame-by-frame behavioral coding. Reliability coding was completed for all trials with 99.99% agreement.

Results and Discussion

Across all experiments, we analyzed children’s search time and switching behavior using separate linear mixed-effects models with trial type and age as predictors and random intercepts for

participants to account for individual variability. Statistical significance was evaluated using a critical value of $t = 2.01$, based on a post hoc sensitivity analysis. Trial order had no effect on children’s search time or switching behavior ($p > .1$) and was therefore not included in subsequent analyses.

We first compared the time that children spent searching in the first location they approached in the opaque versus transparent test trials.² On the transparent trial, children searched significantly longer in the first location they approached ($\beta = 11.76$, $SE = 1.59$), $t(52) = 7.45$, $p < .001$, $d = 1.17$. Age was also a significant predictor of search time ($\beta = -6.02$, $SE = 2.12$), $t(51) = -2.85$, $p < .01$, $d = .54$ (see Figure 2).

We also examined children’s switching behavior. Because children who searched longer in the first search location necessarily had less time to switch, search time and switching were expected to be related. A separate linear mixed-effects model revealed that children made significantly fewer switches between search locations ($\beta = -1.04$, $SE = 0.19$), $t(52) = -5.45$, $p < .001$, $d = .85$, compared to the opaque trial. Age was also a significant predictor of children’s switching behavior ($\beta = 0.90$, $SE = 0.25$), $t(51) = 3.67$, $p < .001$, $d = .73$. These findings provide initial evidence that, as a group, both 3- and 4-year-old children modify their search strategy, depending on whether the outcome is necessary or merely possible.

Experiment 2

To explore the generality of our findings, Experiment 2 replicated Experiment 1 in an adapted version of the task that did not require manual search. This version was completed on a touchscreen tablet and replaced the manual search component with lever taps on gumball machines. Although children produced discrete taps, for consistency across experiments and to control for individual variability in children’s tapping behavior, we again recorded the amount of time spent searching (in seconds), the number of switches between machines, and the first search location that children approached in the necessary (transparent) trial as dependent variables. This experiment was not preregistered.

Method

Participants

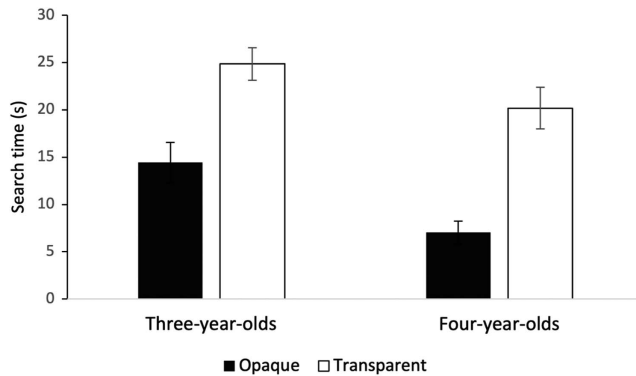
A total of 48 children participated. Children were recruited from the lab’s participant database, preschools, and museums in the broader San Diego area. Ethics approval for this study was granted by University of California, San Diego’s Institutional Review Board. Participants included 24 3-year-olds (11 females, $M_{\text{age}} = 3.5$ years; range = 3;1–3;11) and 24 4-year-olds (13 females, $M_{\text{age}} = 4.5$ years; range = 4;0–4;11). A post hoc sensitivity analysis indicated that we

¹ When we conduct the same analysis without the children who began searching in the incorrect location in the transparent trial ($n = 3$), the findings remain the same: Children spend significantly longer searching the first location in the transparent trial, compared to the opaque trial ($\beta = 12.40$, $SE = 1.63$), $t(49) = 7.62$, $p < .001$, $d = 1.25$, and switch locations significantly fewer times in the transparent trial, compared to the opaque trial ($\beta = -1.02$, $SE = 0.20$), $t(49) = -5.08$, $p < .001$, $d = .82$.

² One 3-year-old was identified as an outlier and excluded because the number of times they switched was more than 3 SDs above the mean for both trials. However, including this child had no impact on our primary outcome of interest (i.e., search time).

Figure 2

Three- and 4-Year-Olds' Search Time in the First Location Approached in Transparent and Opaque Trials in Experiment 1



Note. Error bars represent standard error of the mean.

were well-powered (power = .99) to detect effects at $\alpha = .05$ (critical $t = 2.01$).

Children were either monolingual English speakers or bilingual speakers with English as a primary language. An additional four 3-year-olds and two 4-year-olds were tested but excluded (three withdrew, three technical issues). Thirty-three participants reported racial demographic information, identifying as White (54.5%), Asian (9%), Hispanic or Latino (3%), Native American (3%), and multiracial (30.3%), and 15 families reported their average yearly income, with 11 (73%) families earning more than \$100,000.

Materials

Materials included a 12.9 in. display iPad Pro with a kickstand case, a Bluetooth remote control clicker, a child-size mitten, and stickers.

The study was created in PowerPoint. Search boxes were replaced with images of gumball machines that could be hidden behind a curtain and attached to the openings of fully transparent or fully opaque Y-shaped tubes. Each gumball machine was filled with blue marbles and had a lever that children could tap to dispense the objects inside. Each tap corresponded to a single object being released from the opening on the bottom. Children were instructed to keep searching (tapping the lever) until the target object (a yellow square) came out.

Despite these changes to the framing, the procedure remained largely the same as Experiment 1. Since the procedure was intuitive to children, we reduced the number of familiarization trials from four to two.

Procedure

Children were instructed to find as many yellow squares as they could to win stickers. The screen then showed two gumball machines filled with blue marbles (see Figure 3A). The experimenter demonstrated how the machines worked, and children were given a chance to press each machine's lever once.

Next, a new machine appeared on screen, and children watched as the target object was dropped inside. They were then asked to tap the lever until the yellow square came out. This procedure was repeated again with a new machine, occluded by a black curtain. Two new

occluded machines then appeared on the screen, each behind a curtain, and children were told that a yellow square had already been hidden inside one of the machines (see Figure 3B). During this familiarization trial, the square was hidden inside the left machine and was dispensed on the sixth press. When the child succeeded, they were rewarded with a sticker. Finally, two new occluded machines appeared on the screen, and children were instructed to find the yellow square before the timer ended. This time, the target object was hidden inside the right machine and was dispensed on the third press. As in Experiment 1, if children found the target object before the timer ended, they were awarded another sticker.

The tube familiarization, test trial, and coding scheme were identical to Experiment 1 (see Figure 3C and 3D). Three 3-year-olds and one 4-year-old started searching in the incorrect location on the transparent trial, and these children were included in analyses.³ We used Datavyu to code search times, and reliability coding on 100% of trials resulted in 99.43% agreement.

Results and Discussion

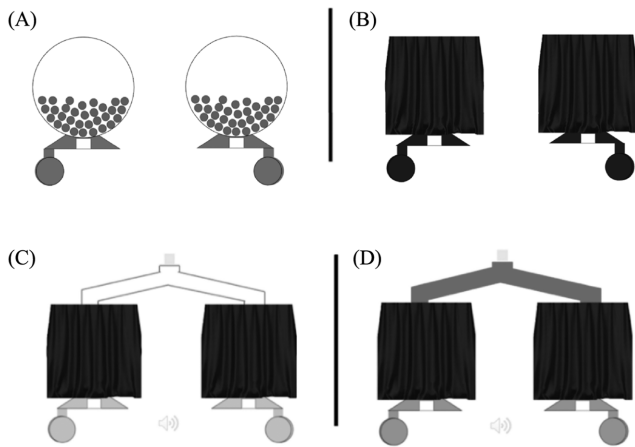
As in Experiment 1, trial order had no effect on children's search time or switching behavior ($p > .1$) and was therefore not included in subsequent analyses. Again, children spent significantly longer searching the first location they approached in the transparent trial ($\beta = 10.03$, $SE = 1.76$), $t(47) = 5.69$, $p < .001$, $d = .88$, compared to the opaque trial. Age did not significantly predict participants' search time ($\beta = -0.09$, $SE = 2.82$), $t(46) = -0.03$, $p = .97$, $d = .01$. However, there was a marginal interaction between trial type and age ($\beta = 6.63$, $SE = 3.42$), $t(46) = 1.94$, $p = .059$, with 4-year-olds showing a more pronounced difference between trial types (see Figure 4).

A separate linear mixed-effects model revealed that children also made significantly fewer switches between search locations ($\beta = -2.10$, $SE = 0.63$), $t(47) = -3.31$, $p < .01$, $d = .40$, during the transparent trial, compared to the opaque trial. Age did not significantly predict participants' switching behavior ($\beta = -0.23$, $SE = 1.39$), $t(46) = -0.17$, $p = .87$, $d = .04$.

Across Experiments 1 and 2, 3- and 4-year-olds differentiated between necessary and merely possible outcomes in their search behavior. However, it is possible that these effects were driven by a difference in the strength of the representation between the two conditions, resulting in greater certainty when the location of the object was known and observable. Specifically, because the transparent tubes allowed children direct visual access to the trajectory of the object, while the opaque set did not, children's representations of the necessary and possible outcomes were based on qualitatively different types of evidence (i.e., direct observation vs. inference). Indeed, prior work suggests that visually tracking an object may lead to a stronger representation (Call & Carpenter, 2001), which could have resulted in the relatively longer search time observed in the transparent trial in Experiments 1 and 2. In Experiment 3, we

³ When we conduct the same analysis without the children who began searching in the incorrect location in the transparent trial ($n = 4$), the findings remain the same: Children spend significantly longer searching the first location in the transparent trial, compared to the opaque trial ($\beta = 11.21$, $SE = 1.81$), $t(43) = 6.19$, $p < .001$, $d = 1.0$, and switch locations significantly fewer times in the transparent trial, compared to the opaque trial ($\beta = -2.27$, $SE = .67$), $t(43) = -3.37$, $p < .01$, $d = .42$.

Figure 3
Experiment 2 Stimuli



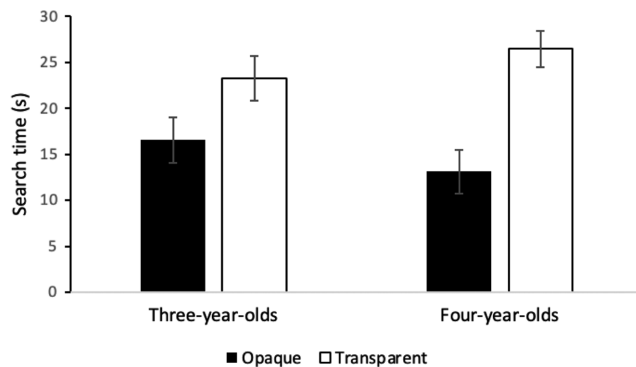
Note. (A) Introductory phase presenting two machines. (B) Example familiarization trial with two occluded machines. (C) Test trial with transparent Y-shaped tubes (necessary trial). (D) Test trial with opaque Y-shaped tubes (merely possible trial).

investigate this alternative explanation using a modified version of Experiment 2 that removes the possibility of object tracking.

Experiment 3

In Experiment 3, we modified the task so that children's representation of the necessary and possible outcomes was both reached indirectly, via inference. The fully transparent set of tubes used in the necessary trial in Experiments 1 and 2 was replaced with a set of tubes with one opaque and one transparent arm. Critically, in this condition, the target object always fell through the opaque branch, requiring children to infer (rather than directly observe) that the object must have fallen through the opaque arm. This design ensures that children's search behavior in both test trials is based on an inference, thereby matching the representation strength across

Figure 4
Three- and 4-Year-Olds' Search Time in the First Location Approached in Transparent and Opaque Trials in Experiment 2



Note. Error bars represent standard error of the mean.

conditions. The method and analysis plan for Experiment 3 was preregistered at <https://aspredicted.org/d9nr-4nf3.pdf>.

Method

Participants

As in Experiment 2, a total of 48 children participated. Children were recruited from the lab's participant database and preschools in the broader San Diego area. Ethics approval for this study was granted by University of California, San Diego's Institutional Review Board. Participants included 24 3-year-olds (10 females, $M_{\text{age}} = 3.5$ years, range = 3;0–3;11) and 24 4-year-olds (12 females, $M_{\text{age}} = 4.49$ years, range = 4;0–4;11). A post hoc sensitivity analysis indicated that primary analyses were well-powered (power = .83) to detect effects at $\alpha = .05$ (critical $t = 2.01$). Children were either monolingual English speakers or bilingual speakers with English as a primary language. An additional eight 3-year-olds and one 4-year-old were tested but excluded (one withdrew, seven technical issues, one inattention during test trials). Forty-one participants reported racial demographic information, identifying as White (56%), Asian (12.1%), Hispanic or Latino (4.8%), multiracial (24.3%), and other (2.4%), and 24 families reported their average yearly income, with 21 (87.5%) families earning more than \$100,000.

Materials

Materials were identical to Experiment 2.

Procedure

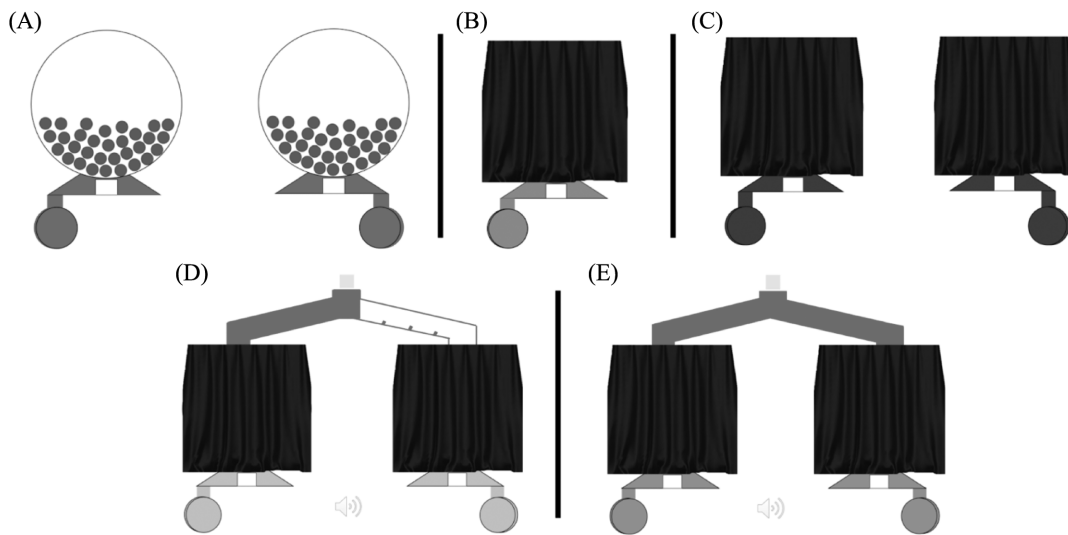
The procedure remained largely the same as Experiment 2. Given the increased task demands in Experiment 3, we increased the number of familiarization trials from two back to four (see Figure 5).

The introductory phase (see Figure 5A and 5B) was similar to Experiment 2. Children were instructed that their job would be to find as many yellow squares as they could to win stickers. The screen then showed two gumball machines filled with blue marbles. The experimenter demonstrated how the machines worked, and children were given a chance to press each machine's lever once. Next, children were shown one machine occluded by a curtain. Children watched as a yellow square was dropped inside and had the opportunity to press the machine's lever to dispense it. Here, the yellow square was dispensed on the fifth press.

Children then observed four familiarization trials (two without a time constraint, two with a time constraint; see Figure 5C) in which two new occluded machines appeared on the screen, and children were told that a yellow square had already been hidden inside one of the machines. We counterbalanced the machine where the square was hidden across trials, and each trial required four to six lever presses to reveal the yellow square.

The tube familiarization procedure was identical to the first two experiments, except that children saw three tubes in the following order: a fully transparent tube, a fully opaque tube, and a half-opaque tube. Objects were dropped in the following pseudorandom orders: L-R-R-L (fully transparent tube), L-R-L-R (fully opaque tube), and R-L-L-R (half-opaque tube). The test trials and coding scheme were the same as Experiments 1 and 2, but the fully transparent tube was replaced with the half-opaque tube.

Figure 5
Experiment 3 Stimuli



Note. (A) Introductory phase presenting two machines. (B) Familiarization with one machine. (C) Example familiarization trial with two occluded machines. (D) Test trial with half-opaque Y-shaped tubes (necessary trial). (E) Test trial with opaque Y-shaped tubes (merely possible trial).

Eight 3-year-olds and 11 4-year-olds started searching in the incorrect location on the transparent trial. We used Datavyu to code search times, and reliability coding on 100% of trials resulted in 97.4% agreement.

Results and Discussion

In line with the previous two experiments, trial order had no effect on children's search behavior ($p > .05$) and was therefore not included in subsequent analyses. As in Experiments 1 and 2, children spent significantly longer searching the first location approached during the half-opaque trial ($\beta = 5.35$, $SE = 1.51$), $t(47) = 3.55$, $p < .001$, $d = .43$, compared to the fully opaque trial. Age did not significantly predict participants' search time ($\beta = -3.93$, $SE = 3.24$), $t(46) = -1.21$, $p = .23$, $d = .31$ (see Figure 6A).

Children also made significantly fewer switches between search locations ($\beta = -1.69$, $SE = 0.56$), $t(47) = -3.01$, $p < .01$, $d = .27$, during the transparent trial, compared to the opaque trial, and age did not significantly predict participants' switching behavior ($\beta = 1.90$, $SE = 1.71$), $t(46) = 1.11$, $p = .27$, $d = .30$. In sum, although this modified task was more challenging than the original, our results generally replicate findings from Experiments 1 and 2 while controlling for object tracking.⁴

Exploratory Analyses

Following Experiments 1 and 2, the planned analyses reported above include all children, regardless of which box they searched first. However, because a sizable number of children in Experiment 3 began their search in the incorrect location in the half-opaque trial, we conducted additional exploratory analyses to better understand this unexpected pattern of search behavior. Specifically, we separately

analyzed the subset of children who searched in the correct location first and the subset who did not.

For the subset of children in both age groups ($n = 16$ 3-year-olds, $n = 13$ 4-year-olds) who began searching in the correct location in the half-opaque (necessary) trial, a linear mixed-effects model revealed that they also spent more time searching in the first location they approached in the half-opaque trial ($\beta = 9.01$, $SE = 2.20$), $t(28) = 4.10$, $p < .001$, $d = .78$, compared to the fully opaque trial (see Figure 6B). A post hoc sensitivity analysis indicated that, despite the smaller sample size, this analysis was also well-powered (power = .98) to detect effects at $\alpha = .05$ (critical $t = 2.05$).

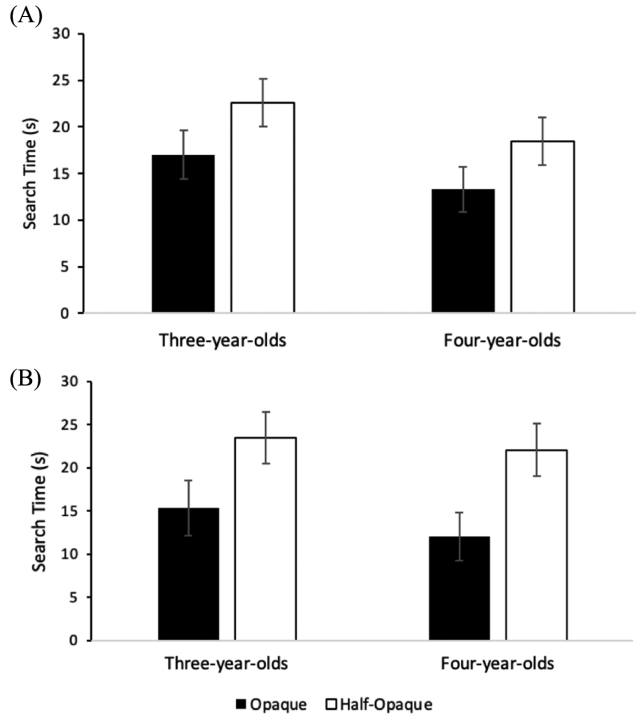
In contrast, the smaller subset of children ($n = 8$ 3-year-olds, $n = 11$ 4-year-olds) who began searching in the incorrect location in the half-opaque (necessary) trial did not show the predicted difference in search behavior. A linear mixed-effects model revealed that children in this subset did not spend significantly more time searching the first location they approached during the half-opaque trial compared to the fully opaque trial ($\beta = -0.22$, $SE = 0.83$), $t(18) = -0.27$, $p = .788$, $d = .02$. However, a post hoc sensitivity analysis indicated that we were underpowered to detect effects at $\alpha = .05$ and power = .80 (critical $t = 2.1$). We note that although this analysis is exploratory and underpowered, the lack of any trend indicating numerical differences in search behavior suggests that this subset of children did not show the same search patterns as children who appropriately began searching in the inferred correct location.

Further exploration of the data reveals that the children in this subset (see Table 1; Other) showed a variety of different search

⁴ This statistical approach deviates from our preregistered plan to analyze 3- and 4-year-olds separately using paired t tests and repeated measures analyses of variance. However, the preregistered analyses produced the same overall pattern of results and supported the same conclusions reported here. Full preregistered analyses can be found at <https://researchbox.org/4711>.

Figure 6

Three- and 4-Year-Olds' Search Time in the First Location Approached in Transparent and Opaque Trials in Experiment 3



Note. The top figure (A) shows the amount of search time spent in the first location for each trial, including all 3- and 4-year-olds in Experiment 3. The bottom figure (B) shows the same data, but excluding 3- and 4-year-olds that began searching in the incorrect location in the half-opaque trial. Error bars represent standard error of the mean.

patterns: Some spent the entire time length they had available searching in the first location they approached in both trials, while others spent a roughly equivalent amount of time in the first location in both trials. Other children spent more time overall in the correct location in the necessary trial, despite the fact that they began

searching in the incorrect location. Since much of the search behavior observed in this smaller group was idiosyncratic, it is difficult to draw clear conclusions. We return to consider possible explanations for this in the General Discussion section.

Transparency and Openness

Materials and data for all experiments are openly available at <https://researchbox.org/4711>. Statistical analyses were conducted in R (Version 4.1.1).

General Discussion

Across three studies, we show that the modal status of an outcome—whether it is necessary or merely possible—systematically regulates how preschoolers allocate their search effort over time. By tracking children's exploration dynamics rather than their discrete choices, we provide a process-level measure of how children respond to outcomes that are necessary or merely possible. In Experiment 1, both 3- and 4-year-olds spent relatively less time searching the first location they approached and switched more frequently when outcomes were merely possible than when they were necessary. In Experiment 2, we replicated this pattern using a touchscreen version of the task. In Experiment 3, we showed that the modulation of search behavior persisted when representational strength was equated across conditions by requiring children to infer the object's trajectory in both trials. Together, these findings indicate that children's exploration is aligned with modal status, indicating that continuous search behavior might provide a more sensitive metric to detect early representations of modal status.

Notably, the modified task used in Experiment 3 was more challenging for all participants, leading a substantial subset of both 3- and 4-year-olds to begin their search in the incorrect location (i.e., the box beneath the transparent arm) in the necessary condition. Whereas some children immediately redirected their search—potentially “double-checking” their inference (see Bohus et al., 2023, for similar behavior when reasoning by exclusion)—others persisted in the incorrect box. This pattern likely reflects the increased processing demands posed by the inferential nature of the task. Unlike Experiments 1 and 2, success in Experiment 3 required

Table 1
Children's Qualitative Search Patterns in Experiments 1, 2, and 3

Experiment	Age	Correct pattern	Incorrect (reverse) pattern	Never switched	Other
Experiment 1	3-year-olds	16	4	6	1
	4-year-olds	20	3	1	2
Experiment 2	3-year-olds	10	2	9	4
	4-year-olds	17	3	3	1
Experiment 3	3-year-olds	8	2	6	8
	4-year-olds	10	1	3	11

Note. The “Correct pattern” column includes children who began searching in the location where they observed the target object in the transparent trial and spent relatively longer in the first location they approached in the transparent (necessary) trial than in the opaque (merely possible) trial. The “Incorrect (reverse) pattern” column includes children who began searching in the correct location in the transparent trial and who spent relatively longer in the first location they approached in the opaque trial than in the transparent trial. The “Never switched” column includes children who began searching in the correct first location in the transparent trial and spent all 31 s searching a single location in both test trials. The “Other” column includes children who began searching in the incorrect location in the transparent trial or who spent an identical amount of time in both trials ($n = 1$; Experiment 2, 3-year-old).

reasoning through a disjunctive inference: (a) The object fell through either the transparent or opaque arm; (b) I did not see it fall through the transparent arm; (c) therefore, it must have fallen through the opaque arm. It also required coordinating this inference with perceptual cues and action planning under time pressure. Prior work shows that such demands can mask underlying inferential abilities unless the task structure is simplified (Alderete & Xu, 2023; Cesana-Arlotti et al., 2018, 2022; Gautam et al., 2021; Grigoroglou et al., 2019; Mody & Carey, 2016). Moreover, in our task, children had to *infer* negation from the absence of visual information (as opposed to directly observing negative evidence, such as being shown an empty cup) while simultaneously reasoning about physical events (i.e., object trajectories, the role of transparency/opacity) and planning an efficient search, creating substantial load on executive and representational resources.

Why, then, did 3-year-olds show clearer modal sensitivity here than in many prior forced-choice tasks (e.g., Leahy, 2023; Leahy et al., 2022; Mody & Carey, 2016)? Our search paradigm was designed to separate our assessment of children's modal representations from how they behave when there are multiple possibilities present. Rather than requiring children to commit to a single option, the task allowed them to distribute their search effort continuously across alternatives. As a result, differences in modal representations could be expressed through graded patterns of exploration rather than through a single discrete response. When outcomes were merely possible, children distributed their search time between the two locations, whereas when outcomes were necessary, they persisted longer in a single location. This suggests that search time provides a more sensitive behavioral index of how children represent modal status. More broadly, measuring exploration dynamics captures how children naturally represent and respond to merely possible outcomes.

Although our analyses focused on group-level patterns, children showed substantial individual variation in how they allocated search. Prior work has suggested that although the majority of 3-year-olds behave in ways consistent with a minimal representation, simulating a single outcome and foreclosing on all others, only a minority (approximately 20%) respond as genuinely modal reasoners (Leahy et al., 2022; Supporting Information). In our data, only a minority of children (33% across all experiments) restricted their search to one location (i.e., persisted) when two locations were possible. Although this behavior is, in principle, compatible with the minimal account, multiple alternative explanations cannot be ruled out: It could arise from perseveration, limited planning, failure to attend to the cued trial start, or difficulty coordinating search under a time constraint. Because our measure indexes how children distribute exploration rather than directly revealing the contents of their representations, it is best suited for detecting process-level sensitivity to modality at the group level. Disentangling the sources of individual variability will require study designs that independently manipulate representational, motivational, and executive factors.

Future studies can build on this process-level approach to examine how individual differences in search behavior reflect variation in how children represent and act on uncertainty in other contexts. For example, future research could consider whether children's sensitivity to necessity versus possibility is stable across domains (e.g., causal, epistemic, and deontic modalities) and also explore the relationship (if any) with related abilities (e.g., response to uncertainty/ambiguity, belief updating, probabilistic reasoning, logical quantification). To better understand the trajectory of

children's sensitivity to modal representations, subsequent studies may also include younger participants or follow children longitudinally to observe whether graded responses in early life predict later explicit reasoning about, for example, counterfactuals and hypothetical alternatives.

Additional work is also needed to better understand the relationship between modal and probabilistic reasoning. As in prior research, our novel design cannot fully distinguish the boundary between these capacities. Indeed, in comparative cognition research, some have argued that apparently logical behavior in nonhuman animals may emerge from deductive processes in which probability distributions are updated in response to new evidence (e.g., "Chrysippus' dog," Rescorla, 2009; chimpanzees, Hanus & Call, 2014). Specifically, after receiving evidence that a reward is not in Location A or B, an individual will restrict their search to Location C via probabilistic updating rather than explicit modal or logical representations. Notably, our results provide evidence against a purely probabilistic interpretation: During the search phase of the current task, negative evidence is matched across conditions (i.e., there is no target object in either search location in both necessary and merely possible trials). If children, like nonhuman animals, were merely updating their probability distribution in light of negative evidence, their search behavior would have been comparable across trials. Of course, it remains possible that probabilistic reasoning played some role in children's behavior earlier in the task—for example, they could have used visual access to the target object (or lack thereof) to inform a probabilistic judgment that later guided their search behavior. Clarifying whether and how probabilistic representations interact with modal reasoning presents an important direction for future research.

In sum, we show that preschoolers' exploration behavior is systematically shaped by whether an outcome is necessary or merely possible. By introducing a continuous, process-based measure of search, we provide a new way to assess how children represent and act on merely possible outcomes. This approach suggests an account of why forced-choice paradigms often underestimate early sensitivity to possibility: Modal concepts may be more readily expressed through graded patterns of exploration rather than binary decisions. More broadly, treating modality as a variable impacting action opens new avenues for studying how children reason, learn, and plan under uncertainty.

References

- Alderete, S., & Xu, F. (2023). Three-year-old children's reasoning about possibilities. *Cognition*, 237, Article 105472. <https://doi.org/10.1016/j.cognition.2023.105472>
- Baumann, L., Schidelko, L. P., Proft, M., & Rakoczy, H. (2024). *Even 3- and 4-year-olds master some modal reasoning tasks—If they have a more agentive structure*. OSF. <https://osf.io/5bkfa/download>
- Beck, S. R., Robinson, E. J., Carroll, D. J., & Apperly, I. A. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development*, 77(2), 413–426. <https://doi.org/10.1111/j.1467-8624.2006.00879.x>
- Begus, K., & Bonawitz, E. (2024). Infants evaluate informativeness of evidence and predict causal events as revealed in theta oscillations and predictive looking. *Communications Psychology*, 2(1), Article 77. <https://doi.org/10.1038/s44271-024-00131-3>
- Begus, K., Gliga, T., & Southgate, V. (2016). Infants' preferences for native speakers are associated with an expectation of information. *Proceedings of the National Academy of Sciences of the United States of America*, 113(44), 12397–12402. <https://doi.org/10.1073/pnas.1603261113>

- Blanco, N. J., & Sloutsky, V. M. (2021). Systematic exploration and uncertainty dominate young children's choices. *Developmental Science*, 24(2), Article e13026. <https://doi.org/10.1111/desc.13026>
- Bohus, K. A., Cesana-Arlotti, N., Martín-Salguero, A., & Bonatti, L. L. (2023). The scope and role of deduction in infant cognition. *Current Biology*, 33(18), 4014–4020.e5. <https://doi.org/10.1016/j.cub.2023.08.028>
- Bonawitz, E. B., van Schijndel, T. J. P., Friel, D., & Schulz, L. (2012). Children balance theories and evidence in exploration, explanation, and learning. *Cognitive Psychology*, 64(4), 215–234. <https://doi.org/10.1016/j.cogpsych.2011.12.002>
- Call, J., & Carpenter, M. (2001). Do apes and children know what they have seen? *Animal Cognition*, 3, 207–220. <https://doi.org/10.1007/s100710100078>
- Cesana-Arlotti, N., Jáuregui, S., Mazalik, P., Nichols, S., & Halberda, J. (2025). Logical concepts of (im)possibility guide young children's decision-making. *Developmental Science*, 28(5), Article e70044. <https://doi.org/10.1111/desc.70044>
- Cesana-Arlotti, N., Kovács, Á. M., & Téglás, E. (2020). Infants recruit logic to learn about the social world. *Nature Communications*, 11(1), Article 5999. <https://doi.org/10.1038/s41467-020-19734-5>
- Cesana-Arlotti, N., Martín, A., Téglás, E., Vorobyova, L., Cetnarski, R., & Bonatti, L. L. (2018). Precursors of logical reasoning in preverbal human infants. *Science*, 359(6381), 1263–1266. <https://doi.org/10.1126/science.aao3539>
- Cesana-Arlotti, N., Varga, B., & Téglás, E. (2022). The pupillometry of the possible: An investigation of infants' representation of alternative possibilities. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 377(1866), Article 20210343. <https://doi.org/10.1098/rstb.2021.0343>
- Cook, C., Goodman, N. D., & Schulz, L. E. (2011). Where science starts: Spontaneous experiments in preschoolers' exploratory play. *Cognition*, 120(3), 341–349. <https://doi.org/10.1016/j.cognition.2011.03.003>
- Dahmani, A., Amir, D., & Gopnik, A. (2025). Beyond rewards: How information value and time horizon shape exploration. *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 47, p. 6637). Cognitive Science Society.
- Denison, S., & Xu, F. (2010). Integrating physical constraints in statistical inference by 11-month-old infants. *Cognitive Science*, 34(5), 885–908. <https://doi.org/10.1111/j.1551-6709.2010.01111.x>
- Denison, S., & Xu, F. (2014). The origins of probabilistic inference in human infants. *Cognition*, 130(3), 335–347. <https://doi.org/10.1016/j.cognition.2013.12.001>
- Denison, S., & Xu, F. (2019). Infant statisticians: The origins of reasoning under uncertainty. *Perspectives on Psychological Science*, 14(4), 499–509. <https://doi.org/10.1177/1745691619847201>
- Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: Evidence from infants' manual search. *Developmental Science*, 6(5), 568–584. <https://doi.org/10.1111/1467-7687.00313>
- Fodor, J. A. (1975). *The language of thought*. Harvard University Press. <https://doi.org/10.2307/2184356>
- Gautam, S., Suddendorf, T., & Redshaw, J. (2021). When can young children reason about an exclusive disjunction? A follow up to. *Cognition*, 207, Article 104507. <https://doi.org/10.1016/j.cognition.2020.104507>
- Goddu, M. K., & Gopnik, A. (2024). The development of human causal learning and reasoning. *Nature Reviews Psychology*, 3(5), 319–339. <https://doi.org/10.1038/s44159-024-00300-5>
- Goddu, M. K., Sullivan, J. N., & Walker, C. M. (2021). Toddlers learn and flexibly apply multiple possibilities. *Child Development*, 92(6), 2244–2251. <https://doi.org/10.1111/cdev.13668>
- Gopnik, A. (2000). Explanation as orgasm and the drive for causal knowledge: The function, evolution, and phenomenology of the theory formation system. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 299–326). The MIT Press. <https://doi.org/10.7551/mitpress/2930.003.0018>
- Gopnik, A. (2024, December). *Empowerment gain as causal learning, causal learning as empowerment gain: A bridge between Bayesian causal hypothesis testing and reinforcement learning* [Paper presentation]. Presented at the 29th Meeting of The Philosophy of Science Association, New Orleans.
- Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin*, 138(6), 1085–1108. <https://doi.org/10.1037/a0028044>
- Goupil, L., Romand-Monnier, M., & Kouider, S. (2016). Infants ask for help when they know they don't know. *Proceedings of the National Academy of Sciences of the United States of America*, 113(13), 3492–3496. <https://doi.org/10.1073/pnas.1515129113>
- Grigoroglou, M., Chan, S., & Ganea, P. A. (2019). Toddlers' understanding and use of verbal negation in inferential reasoning search tasks. *Journal of Experimental Child Psychology*, 183, 222–241. <https://doi.org/10.1016/j.jecp.2019.02.004>
- Hanus, D., & Call, J. (2014). When maths trumps logic: Probabilistic judgements in chimpanzees. *Biology Letters*, 10(12), Article 20140892. <https://doi.org/10.1098/rsbl.2014.0892>
- Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences*, 8(6), 280–285. <https://doi.org/10.1016/j.tics.2004.04.001>
- Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proceedings of the National Academy of Sciences of the United States of America*, 107(43), 18243–18250. <https://doi.org/10.1073/pnas.1012933107>
- Johnson-Laird, P. N., Byrne, R. M., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review*, 99(3), 418–439. <https://doi.org/10.1037/0033-295X.99.3.418>
- Kahneman, D., & Tversky, A. (1981). *The simulation heuristic*. Stanford University Press. <https://doi.org/10.1017/CBO9780511809477.015>
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLOS ONE*, 7(5), Article e36399. <https://doi.org/10.1371/journal.pone.0036399>
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks effect in infant auditory attention. *Child Development*, 85(5), 1795–1804. <https://doi.org/10.1111/cdev.12263>
- Kim, S., Paulus, M., Sodian, B., & Proust, J. (2016). Young children's sensitivity to their own ignorance in informing others. *PLOS ONE*, 11(3), Article e0152595. <https://doi.org/10.1371/journal.pone.0152595>
- Lapidow, E., & Bonawitz, E. (2023). What's in the box? Preschoolers consider ambiguity, expected value, and information for future decisions in explore-exploit tasks. *Open Mind*, 7, 855–878. https://doi.org/10.1162/opmi_a_00110
- Lapidow, E., Killeen, I., & Walker, C. M. (2022). Learning to recognize uncertainty vs. recognizing uncertainty to learn: Confidence judgments and exploration decisions in preschoolers. *Developmental Science*, 25(2), Article e13178. <https://doi.org/10.1111/desc.13178>
- Lapidow, E., & Walker, C. M. (2020). Informative experimentation in intuitive science: Children select and learn from their own causal interventions. *Cognition*, 201, Article 104315. <https://doi.org/10.1016/j.cognition.2020.104315>
- Leahy, B. (2023). Don't you see the possibilities? Young preschoolers may lack possibility concepts. *Developmental Science*, 26(6), Article e13400. <https://doi.org/10.1111/desc.13400>
- Leahy, B. (2024). Many preschoolers do not distinguish the possible from the impossible in a marble-catching task. *Journal of Experimental Child Psychology*, 238, Article 105794. <https://doi.org/10.1016/j.jecp.2023.105794>
- Leahy, B., Huemer, M., Steele, M., Alderete, S., & Carey, S. (2022). Minimal representations of possibility at age 3. *Proceedings of the National Academy of Sciences of the United States of America*, 119(52), Article e2207499119. <https://doi.org/10.1073/pnas.2207499119>
- Leahy, B., & Zalnieriunas, E. (2021). Might and might not: Children's conceptual development and the acquisition of modal verbs. *Semantics and Linguistic Theory* (pp. 426–445).

- Leahy, B. P., & Carey, S. E. (2020). The acquisition of modal concepts. *Trends in Cognitive Sciences*, 24(1), 65–78. <https://doi.org/10.1016/j.tics.2019.11.004>
- Liquin, E. G., & Gopnik, A. (2022). Children are more exploratory and learn more than adults in an approach-avoid task. *Cognition*, 218, Article 104940. <https://doi.org/10.1016/j.cognition.2021.104940>
- Liquin, E. G., & Lombrozo, T. (2020). Explanation-seeking curiosity in childhood. *Current Opinion in Behavioral Sciences*, 35, 14–20. <https://doi.org/10.1016/j.cobeha.2020.05.012>
- Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences*, 10(10), 464–470. <https://doi.org/10.1016/j.tics.2006.08.004>
- Mody, S., & Carey, S. (2016). The emergence of reasoning by the disjunctive syllogism in early childhood. *Cognition*, 154, 40–48. <https://doi.org/10.1016/j.cognition.2016.05.012>
- Ozturk, O., & Papafragou, A. (2015). The acquisition of epistemic modality: From semantic meaning to pragmatic interpretation. *Language Learning and Development*, 11(3), 191–214. <https://doi.org/10.1080/15475441.2014.905169>
- Perez, J., & Feigenson, L. (2022). Violations of expectation trigger infants to search for explanations. *Cognition*, 218, Article 104942. <https://doi.org/10.1016/j.cognition.2021.104942>
- Phillips, J., & Kratzer, A. (2024). Decomposing modal thought. *Psychological Review*, 131(4), 966–992. <https://doi.org/10.1037/rev0000481>
- Redshaw, J., & Suddendorf, T. (2016). Children's and apes' preparatory responses to two mutually exclusive possibilities. *Current Biology*, 26(13), 1758–1762. <https://doi.org/10.1016/j.cub.2016.04.062>
- Rescorla, M. (2009). Chrysippus' dog as a case study in non-linguistic cognition. In R. W. Lurz (Ed.), *The philosophy of animal minds* (pp. 52–71). Cambridge University Press. <https://doi.org/10.1017/cbo9780511819001.004>
- Robinson, E. J., Rowley, M. G., Beck, S. R., Carroll, D. J., & Apperly, I. A. (2006). Children's sensitivity to their own relative ignorance: Handling of possibilities under epistemic and physical uncertainty. *Child Development*, 77(6), 1642–1655. <https://doi.org/10.1111/j.1467-8624.2006.00964.x>
- Rohwer, M., Kloof, D., & Perner, J. (2012). Escape from metaignorance: How children develop an understanding of their own lack of knowledge. *Child Development*, 83(6), 1869–1883. <https://doi.org/10.1111/j.1467-8624.2012.01830.x>
- 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. [https://doi.org/10.1016/s0065-2407\(08\)60348-1](https://doi.org/10.1016/s0065-2407(08)60348-1)
- Ruggeri, A., Sim, Z. L., & Xu, F. (2017). “Why is Toma late to school again?” Preschoolers identify the most informative questions. *Developmental Psychology*, 53(9), 1620–1632. <https://doi.org/10.1037/dev0000340>
- Rule, J. S., Piantadosi, S. T., Cropper, A., Ellis, K., Nye, M., & Tenenbaum, J. B. (2024). Symbolic metaprogram search improves learning efficiency and explains rule learning in humans. *Nature Communications*, 15(1), Article 6847. <https://doi.org/10.1038/s41467-024-50966-x>
- Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, 43(4), 1045–1050. <https://doi.org/10.1037/0012-1649.43.4.1045>
- 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), 642–651. <https://doi.org/10.1037/dev0000278>
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99–118. <https://doi.org/10.2307/1884852>
- Stahl, A. E., & Feigenson, L. (2015). Cognitive development. Observing the unexpected enhances infants' learning and exploration. *Science*, 348(6230), 91–94. <https://doi.org/10.1126/science.aaa3799>
- Stahl, A. E., & Feigenson, L. (2024). Young children distinguish the impossible from the merely improbable. *Proceedings of the National Academy of Sciences of the United States of America*, 121(46), Article e2411297121. <https://doi.org/10.1073/pnas.2411297121>
- Sumner, E., Steyvers, M., & Sarnecka, B. W. (2019). It's not the treasure, it's the hunt: Children are more explorative on an explore/exploit task than adults. In A. K. Goel, C. M. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st annual conference of the Cognitive Science Society* (pp. 2891–2897). Cognitive Science Society.
- Téglás, E., & Bonatti, L. L. (2016). Infants anticipate probabilistic but not deterministic outcomes. *Cognition*, 157, 227–236. <https://doi.org/10.1016/j.cognition.2016.09.003>
- Téglás, E., Girotto, V., Gonzalez, M., & Bonatti, L. L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proceedings of the National Academy of Sciences of the United States of America*, 104(48), 19156–19159. <https://doi.org/10.1073/pnas.0700271104>
- Turan-Küçük, E. N., & Kibbe, M. M. (2024). Three-year-olds' ability to plan for mutually exclusive future possibilities is limited primarily by their representations of possible plans, not possible events. *Cognition*, 244, Article 105712. <https://doi.org/10.1016/j.cognition.2023.105712>
- Turan-Küçük, E. N., & Kibbe, M. M. (2025). Three- and four-year-old children represent mutually exclusive possible identities. *Journal of Experimental Child Psychology*, 249, Article 106078. <https://doi.org/10.1016/j.jecp.2024.106078>
- Van de Walle, G. A., Carey, S., & Prevor, M. (2000). Bases for object individuation in infancy: Evidence from manual search. *Journal of Cognition and Development*, 1(3), 249–280. https://doi.org/10.1207/S15327647JCD0103_1
- Wang, J. J., Yang, Y., Macias, C., & Bonawitz, E. (2021). Children with more uncertainty in their intuitive theories seek domain-relevant information. *Psychological Science*, 32(7), 1147–1156. <https://doi.org/10.1177/0956797621994230>

Received March 4, 2026

Revision received July 28, 2026

Accepted August 5, 2026 ■