UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Possible identities versus possible absence in three-year-olds' modal reasoning

Permalink

https://escholarship.org/uc/item/63f5h72v

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Meade, Madeline G
Braswell, Nathaniel
Santos, Laurie
et al.

Publication Date

2026

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Possible identities versus possible absence in three-year-olds' modal reasoning

Madeline G. Meade (mmeade@ucsd.edu)[1], Nathaniel Braswell[2], Laurie Santos[2], Nicolò Cesana-Arlotti[2] & Federico Rossano[1]

[1]Department of Cognitive Science, University of California, San Diego
[2]Department of Psychology, Yale University

Abstract

Modal concepts—representations of possibility, necessity, and impossibility—play a central role in human thought. Some argue that these are uniquely human and dependent on modal language, whereas others propose earlier developmental and evolutionary emergences. Recent success by rhesus macaques on a modified 3-cup task supports the latter view. Here, three-year-old children (N =144), who struggle with modal words, completed equivalent modified 3-cup tasks across three experiments. When preserving most features of the original 3-cup task—contrasting a guaranteed reward with a possible reward or its absence—children mirrored macaque performance and did not exceed chance. However, children successfully preferred a guaranteed reward to a possible reward or non-reward. A follow-up experiment confirmed that our paradigms did not bias participants towards the guaranteed side. Although this pattern broadly paralleled macaques, evidence for a between-experiments difference in performance was weaker, likely reflecting different primary constraints on 3-year-olds' modal reasoning abilities.

Keywords: modal reasoning; logic; cognitive development

Introduction

The ability to represent and evaluate possibilities pervades and underlies much of human reasoning across domains of cognition (Johnson-Laird et al., 2024). Specifically, marking potential outcomes as either guaranteed, possible, or impossible allows one to quickly differentiate optimal behavior from suboptimal: a guaranteed reward is always preferable to a possible one, a possible reward (regardless of its specific probability of occurring) is preferable to an impossible one, it is far better for an undesirable outcome to be an impossibility than a possibility, etc. However, despite the ubiquity and importance of this ability across cognition (Roese & Olson, 1993; Samet, 1996; Van Hoeck et al., 2014), the origins of this capacity on both evolutionary and developmental scales remain hotly contested.

The most common method of testing for modal competency is a suite of paradigms known as "cup tasks", in which a subject is offered a choice between cups containing rewards of different modal statuses (Mody & Carey, 2016). Specifically, the 3-cup task offers children the choice between a single cup that definitively receives a reward and two other paired cups into which one reward is ambiguously placed. Here, an agent with an understanding of the boundaries between the modal concepts of certainty and possibility, and a desire for the reward, should always choose to first search the single cup.

The clearest current evidence indicates that these capacities develop in children by the ages of 4 to 5, allowing them to accurately prepare for mutually exclusive possible outcomes and make choices that give them the best chance of reward (Leahy, 2024; Leahy et al., 2022; Leahy & Carey, 2020; Redshaw & Suddendorf, 2016). This account is in part supported because children begin comprehending modal language (words that denote possibilities, such as might, must, have to, etc.) by age 4-5; some scholars credit modal language as necessary for scaffolding the comprehension of modal concepts (Leahy & Žalnieriunas, 2021).

In support of this hypothesis, the performance of children younger than 4-to-5-years old on modal tasks is characterized by ambiguity. Specifically, three-year-old children tend to select a guaranteed reward over a possible one in 3-cup-style tasks approximately 60% of the time (Grigoroglou, et al., 2019; Leahy et al., 2022; Leahy, 2024; Mody & Carey, 2016). This percentage constitutes a success when chance level is assigned at 33%, due to the three possible cups one could select, but does not reach ceiling-level performance.

According to one influential proposal, known as the minimal representation account, this result pattern can be cleanly explained by participants simulating only one possible outcome at a time and treating that simulation as ground truth (Leahy & Carey, 2020). Under this account, a minimal representer might guess which of the two paired cups received the reward, and treat that guess as just as guaranteed of a choice as the single-cup side's truly guaranteed reward.

However, three-year-olds have also been found to succeed at conceptually similar modal reasoning tasks (Alderete & Xu, 2023; Andreuccioli et al., 2024; Cesana-Arlotti et al., 2025, Cesana-Arlotti & Halberda, 2024; Turan-Küçük & Kibbe, 2025). Additionally, eye tracking and pupillometric methods have revealed surprisingly complex levels of logical competence even in human infants (Cesana-Arlotti et al., 2018, 2020), who are almost entirely prelinguistic, leaving the door open for a more basal, potentially evolutionarily ancient capacity for modal reasoning.

The clearest way to explore such a possibility is to search for modal competence outside of the human species. Unfortunately, whether nonhuman primates (hereafter primates) can reason about multiple incompatible

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society* ©2026 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

possibilities has also long remained unresolved. Previously, the most direct test of the traditional 3-cup task examined sanctuary-housed chimpanzees; this population chose the single cup holding a certain reward only 50% of the time — in line with performance predictions according to the minimal representation hypothesis —despite succeeding on more complex reasoning tasks (Engelmann, Haux, et al., 2023). Additional studies assessing primates' ability to prepare for multiple mutually exclusive outcomes simultaneously yielded inconsistent results (failures: Redshaw & Suddendorf, 2016; successes: Engelmann et al., 2021; Engelmann, Völter, et al., 2023; Ferrigno et al., 2021; but see Gautam et al., 2021 and Redshaw & Suddendorf, 2024 for alternative interpretations of these results).

More recently, rhesus macaques successfully distinguished guaranteed from merely possible outcomes in a one-shot, modified 3-cup task (Braswell, Meade, et al., 2025). Subjects preferred a cup that contained a guaranteed food reward over one that either held an identical food reward or a non-rewarding item. This performance also matched a pure object location recall task; this contrasts with the minimal representation of possibility account, which predicts above chance—but below-ceiling—performance in a population of minimal representers. Despite this, the same population failed the more traditional version of the 3-cup task, choosing at random between a cup that either was empty or contained food and a cup that definitely contained food.

Why might macaques have succeeded on one experiment but not the other, and what does this modified paradigm have in common with other examples of successes in the developmental field? It may be that representing nothing, in the form of an empty location, as an alternative within a possibilistic representation is more cognitively challenging than reasoning about the possible identity of an item. Thus, typical 3-cup tasks featuring only two items across three possible locations may introduce an additional burden on top of the possibility reasoning itself. This issue may have been enough to obscure the abilities of both primates and young children in many modal reasoning tasks, as tasks involving tracking possible object trajectories (*e.g.*, Redshaw & Suddendorf, 2016) also encounter the same confound of reasoning about only one object over possible locations.

A discrepancy between reasoning about a known item's possible identities and the presence or absence of an item at a given location may also explain the mixed pattern of successes and failures in the developmental area: both children and infants specifically perform well in the tasks involving possible identities (*e.g.*, Alderete & Xu, 2023; Cesana-Arlotti et al., 2018; Turan-Küçük & Kibbe, 2025). Such results may also align with recent research into adult vision science, which found evidence that object files can incorporate information about objects' possible identities, but not multiple possible locations for one object (Brody et al., 2024).

Thus, the question remains: does the pattern of macaque success and failure represent how all primate minds lacking modal vocabulary handle modal challenges, or rather indicate a species-level difference in how nonhuman minds solve them compared to humans? We aimed to extend the aforementioned work with macaques into the developmental domain, specifically examining whether the types of possibilities represented (either an object's unknown identity or its presence or absence in a location) influenced performance in an otherwise identical task. We chose to test three-year-old children, who already have basic verbal abilities, executive function skills, and a history of being tested on 3-cup tasks. Importantly, prior evidence indicates that only a small portion of four-year-olds consistently comprehend modal terms (Leahy & Žalnieriunas, 2021). Thus, this age group allows us to examine modal reasoning in the absence of reliable modal language comprehension.

Current Study

Across three experiments, 144 three-year-old children were recruited from museums and aquariums. In Experiment 1, we tested a more traditional form of the 3-cup task, asking whether three-year-olds would choose a cup with a guaranteed reward over a cup that either contained a reward or contained nothing. In Experiment 2, we deployed the same novel variation of the 3-cup task to which rhesus macaques were exposed; we tested three-year-olds' preferences for a guaranteed reward over a possible reward or a *non-reward*, examining whether the introduction of the non-rewarding item led to improved performance in the otherwise identical task. Finally, Experiment 3 examined children's preferences for a possible reward or non-reward over a guaranteed non-reward, while controlling for non-modal heuristic strategies that could explain success in the modified paradigm.

Experiment 1

Methods

Participants 48 three-year-old children (mean age = 3.44, range = 3.04-3.98, 24 females) were recruited at a museum in the San Diego metropolitan area. The sample size was chosen to be double that of prior studies' sample of three-year-olds (e.g., Mody & Carey, 2016). One additional child was tested but excluded from the analysis due to experimenter error during the demonstration.

Procedure and Apparatus Three identical black boxes made from foam board panels and black tape (~12 cm × 12 cm) served as the cups for the task and were identical to those used with rhesus macaques. Each was open at the top and could pivot open to show its (empty) contents at the start. In line with prior 3-cup experiments with children, identical stickers served as the rewards for the task.

The experimental procedure began with the experimenter showing the participant that all three cups were empty, then pivoting them closed. Next, the two stickers were hidden in the three cups. One sticker was hidden on the single-cup side, matched with the words "Look! I have a sticker! I'm putting it here". The other sticker was ambiguously hidden in the

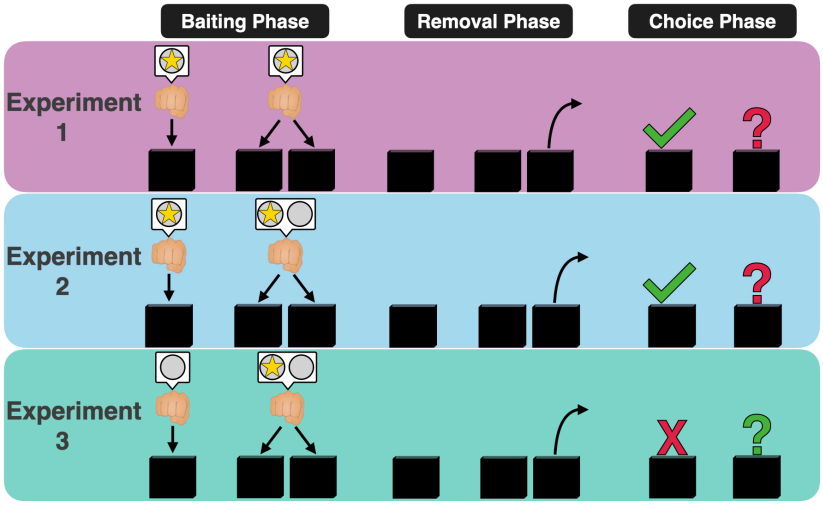

Figure 1: Depiction of all three experiments. Correct choices in the Choice Phase are indicated in green, and reward status is indicated by check mark (guaranteed sticker reward), X (guaranteed paper non-reward), or question mark (potential sticker reward). The side that the single cup appeared on, the order of hiding, and which of the paired cups was removed were counterbalanced in all experiments.

paired-cup side. To accomplish this, the experimenter sequentially dipped a closed fist holding the sticker into both cups, such that the participant was unaware of which cup in the pair received the reward; this was temporally matched with the words, "Look! I have another sticker. I'm putting it either in here, or in here". The experimenter then removed one of the cups from the pair. After doing so, she requested that the participant choose one of the two remaining cups to look in to find a sticker.

The participant's choice was recorded as the first box they interacted with—coded as the first box they verbally requested or physically indicated (by pointing to, grabbing, or looking inside of) upon completion of the demonstration. If the participant failed to make a choice within 30 seconds of the demonstration concluding, they were excluded from analyses. As this was a one-shot design, participants completed only the single trial and immediately received feedback on their decision.

Design We employed eight unique conditions to counterbalance the procedure; we counterbalanced which side of the table the paired-cup side appeared on (left or right), the order of object hiding (single-cup first or paired-cups first), and which of the paired-cups was removed (left or right). The script of the experiment was also updated to match the condition that the participant was assigned to. These conditions were counterbalanced between subjects, balanced across gender, and collapsed for the purpose of our analysis.

Results & Discussion

All analyses were conducted in R (version 4.4.2). In the preregistered analysis, an exact binomial test found that children did not reliably choose the target option above chance (M_{Single} = 60.4%; chance level = 50%, p = .097, 95% CI [0.48, 1.00]). Additionally, exploratory generalized linear models with a binomial link revealed that there were no significant effects of age, sex, or conditions on choice.

Experiment 1's results thus indicate that, when a possible reward is either present or absent on the paired-cup side, we find traditional levels of success for this age group in the 3-cup task (*e.g.,* Mody & Carey, 2016). However, as this task's choice phase offered three-year-olds a single choice between two cups, rather than three, the equivalent performance to previous studies was *not* sufficient to exceed chance in this experiment.

When rhesus macaques were exposed to this version of the 3-cup task, their performance as a group also did not exceed chance (Braswell, Meade, et al., 2025). Thus, the question remained: when given a modified task instead requiring reasoning about the possible identity of the object within a cup, rather than the potential reward or absence of a reward within it, would three-year-olds' performance on the task improve?

Experiment 2

Methods

Participants 48 additional three-year-old children (mean age = 3.53, range = 3.03-3.99, 24 females) were recruited from a museum in the San Diego metropolitan area. One additional child was tested but excluded from the analysis for failing to make a choice within the 30 second decision period.

Procedure and Apparatus Three identical black boxes made from foam board panels and black tape (~12 cm × 12 cm) served as the cups for the task and were identical to those used in Experiment 1. The plain back paper of a sticker was chosen as a neutral, non-rewarding item for the task. Not only

were these papers thus physically matched in size to the stickers, but they were also an item that participants would receive in conjunction with the sticker. As such, we did not expect for participants to find these non-rewards aversive, but merely expected them to be treated as less desirable than the sticker reward.

The experimental procedure was nearly identical to that of Experiment 1, barring the addition of the non-reward to the paired-cup side. Once again, the experimenter showed the participant that all three cups were empty and pivoted them closed. Next, two rewards *and one non-reward* were each hidden in one of the three cups. One sticker was hidden on the single-cup side, again matched with the words "Look! I have a sticker! I'm putting it here". The other sticker and the paper were then ambiguously hidden in the paired-cup side. The experimenter sequentially dipped a closed fist holding both objects into both cups, such that the participant was unaware of which cup in the pair received the reward and which received the non-reward; this was matched with the words, "Look! I have a sticker and a piece of paper. I'm putting one in here, and one in here". The experimenter removed one of the cups from the pair and then allowed the participant to choose one of the two remaining cups to look in to find a sticker, leaving the participant to choose between a cup definitely containing a reward and a cup containing either a reward or a non-reward.

All choice criteria were identical to that of Experiment 1.

Design We used the same eight unique conditions to counterbalance the procedure (balancing the side of the table that the paired-cup side was on, whether the paired-cups or the single-cup was baited first, and which of the paired-cups was removed). Conditions were again counterbalanced between subjects, balanced across gender, and collapsed for the purpose of our analysis.

Results & Discussion

All analyses were again conducted in R (version 4.4.2). In the preregistered analysis, an exact binomial test revealed that three-year-old children chose the single cup with the guaranteed reward over the remaining paired-cup with the possible reward or non-reward above chance levels ($n = 48$, $M_{Single} = 68.75\%$; chance level = 50%, $p = .007$, 95% CI [0.56, 1.00]). A Bayesian contingency-table analysis comparing Experiments 1 and 2 yielded anecdotal-to-moderate evidence in favor of the null hypothesis of equal performance across experiments ($BF_{10} = 0.37$), indicating no reliable difference in choice distributions.

An exploratory generalized linear model with a binomial link revealed no significant effects of participant age, sex, or condition on choice accuracy, although marginal trends were observed for the side of the single cup ($p = .058$)—with children succeeding marginally more when the single cup was on their right—and hiding order ($p = .076$)—succeeding marginally more when the reward was hidden in the single cup first.

When Experiment 2's choice phase offered three-year-olds a single choice between two cups of different possible reward statuses, their performance as a group *did* exceed chance (unlike in Experiment 1). Here as well, we document a similar reasoning pattern to that of rhesus macaques—success when reasoning about objects' identities, but not when considering possible object presence or absence. However, the Bayesian analysis adds a crucial caveat to this result. Three-year-olds were approximately 8% more successful as a group in Experiment 2 compared to Experiment 1; this result was not sufficient evidence to detect a meaningful between-experiment difference in performance. Additionally, this was a far smaller improvement than the >30% increase in success that the rhesus macaques exhibited in an equivalent paradigm (Braswell, Meade, et al., 2025); such a jump in success reflected chance performance in the first experiment, and ceiling-level performance (in reference to a pure memory task) in the other.

Although children were above chance at this task, the alterations that we made to the traditional 3-cup task opened other possible avenues for successful or boosted performance—without the need to reason modally. For example, the introduction of a non-reward into the paradigm could prompt a participant to (non-logically) choose the correct, single-cup side simply from a desire to avoid the side with the non-reward. Additionally, removing one of the cups from the paired-cup side may have disturbed the child's representation of that side, causing them to avoid it. This account, too, would not require a participant to represent the certainty of the single cup when choosing it. Though the second concern would have applied equally to Experiments 1 and 2 and thus could not explain a between-experiment difference, it still poses a possible confound for comparing our study's results to the broader literature on the topic. Therefore, we developed Experiment 3 to control for such accounts.

Experiment 3

Methods

Participants 48 additional three-year-old children (mean age = 3.53, range = 3.03-3.99, 24 females) were recruited from museums and aquariums in the San Diego metropolitan area. One additional child was tested but excluded from the analysis for opting out of the experiment after the demonstration concluded.

Procedure and Apparatus All materials, rewards, and non-rewards were identical to Experiments 1 and 2. Instead of the two sticker rewards and one sticker paper back used in Experiment 1, this design used two non-rewarding sticker backs and only one sticker reward.

The structure of the experimental procedure was also identical to Experiments 1 and 2. However, here a sticker back was hidden on the single-cup side (rather than a reward); the experimenter's script was updated to reflect this. The baiting of the paired-cup side was identical to the procedure

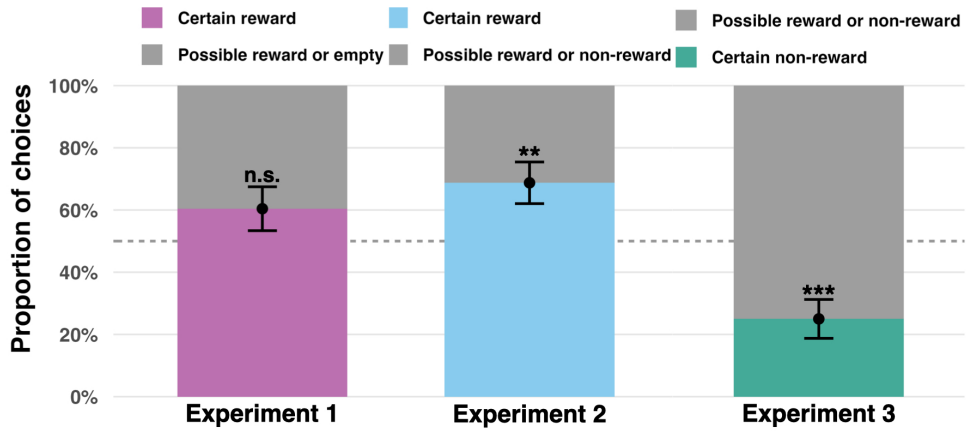

Figure 2: Group-level performance on all experiments, with error bars indicating ±1 standard error of the mean. All certain outcomes are indicated in color and possible outcomes in gray, though the identities of the certain and possible items varied across experiments.

of Experiment 2. Thus, rather than offering a guaranteed reward compared to a possible reward or non-reward as in Experiment 2, Experiment 3 offered participants a possible reward or non-reward in contrast with a guaranteed non-reward during the decision period.

All choice criteria were identical to that of Experiments 1 and 2.

Design All counterbalancing conditions were identical to the previous experiments.

Results & Discussion

Analyses were again conducted in R (version 4.4.2). In the preregistered analysis, an exact binomial test found that three-year-old children preferred the cup containing either the reward or the non-reward over the cup with the certain non-reward (M_{Double} = 75%; chance level = 50%, $p < .001$, 95% CI [0.627, 1.00]). Exploratory analyses utilizing a Bayesian contingency-table analysis comparing Experiments 2 and 3 yielded weak evidence in favor of the null hypothesis of equal performance across experiments (BF_{10} = 0.35), indicating no reliable difference in choice distributions. Furthermore, a generalized linear model with a binomial link revealed no reliable effects of age or sex on choice accuracy; however, there was a significant effect of the side that the single cup appeared on, with higher accuracy when the single cup appeared on the child's left ($p = .017$)—the reverse direction of the Experiment 1's marginal trend—while all other task-structure predictors were non-significant.

Thus, Experiment 3 demonstrated that children could choose a cup containing a possible reward over one that certainly contained a non-reward, and that they did so at rates similar to the success rate of Experiment 2. In designing a study in which a successful reasoner should always choose the paired-cup side, we also received confirmation that no aspects of the modified 3-cup procedure inadvertently biased participants away from the double-cup side.

General Discussion

In Experiment 1, three-year-olds failed to choose a box with a guaranteed reward over one that either contained a reward or was empty, choosing the guaranteed option approximately 60% of the time. In Experiment 2, three-year-olds chose a cup with a guaranteed reward over one that held either a reward or a non-reward approximately 69% of the time, exceeding chance level; however, a between-experiments comparison indicated that these results may not meaningfully differ from those of Experiment 1. Finally, in Experiment 3, three-year-olds also exceeded chance at choosing a cup that either contained a reward or a non-reward over one that certainly contained a non-reward 75% of the time.

Overall, rhesus macaques showed a similar pattern of behavior on the equivalent tasks: failure in our more traditional version of the 3-cup task involving possibly empty locations (succeeding in approximately 48% of the trials), success at the modified version involving possible identities (succeeding in 80% of trials), and success in the control experiment (approximately 81% of trials) (Braswell, Meade, et al., 2025). However, those between-experiment differences were much larger than what was found in this study. Thus, children's performance appeared less strongly modulated by a difficulty in representing possibly empty locations in their decision-making, as reflected in the smaller separation between experiments. While more research is needed, this pattern may reflect a difference of degree rather than kind across species: reasoning about certain types of possibilities may impose only a modest additional computational burden for young children, whereas it may constitute a substantial barrier for nonhuman primates. In particular, reasoning about identity-based possibilities may be more naturally supported by the primate visual system's representations of object files, whereas reasoning about possible locations may not be facilitated (Brody et al., 2024).

Our experimental paradigms differed in several ways from the classic 3-cup task: trials were one-shot decisions that did not allow for the opportunity to improve over subsequent

trials, involved no explicit training on modal concepts or the specific materials of the study, and, most drastically, offered three-year-olds a choice between two boxes when making their decision, not all three. Thus, it is especially striking that Experiment 1's results precisely replicated prior success rates in three-year-olds, despite the different assignment of chance level. It should be noted that even with all three cups present in the decision period, a case can be made for assigning the chance level of a three cup task to 50%; Engelmann, Haux & colleagues (2023), recognizing that another nonlogical path to success is to treat both sides—rather than all three cups—as equivalent and choose randomly between them, analyzed their 3-cup task data according to this stricter criterion.

What other factors might account for differences in performance on our task between primates and children? One possible explanation for the differences in overall success concerns the different strengths of motivation that the two populations may possess. A free ranging primate species may be more intrinsically motivated to obtain a high-value, seasonal reward such as a fruit due to resource competition pressure than a child is to win a sticker. Furthermore, the risk of approaching a human experimenter and stealing from them may provide further incentive to make a correct first decision. Conversely, children in this study were informed of their ability to make only a single choice, but a verbal instruction may not have been as salient to a participant.

Some might be concerned that performance could be driven by avoidance of non-reward options rather than genuine reasoning about possibility. We chose the back paper of a sticker as our non-rewarding item because it was neutrally-valenced, in an attempt to avoid potential confounds of succeeding due to an aversion to the item; however, increasing the stakes (i.e., contrasting a positive reward—a cupcake—with an aversive hot pepper) of a third party character has led to success in one recent modal reasoning paradigm (Cesana-Arlotti et al., 2025).

Another possible confound may be that the baiting procedure itself introduces asymmetries between experiments, since the experimenter's actions were similar regardless of whether one or two items were hidden (and thus were less efficient when hiding one item). However, there was no evidence that counterbalancing affected children's choices, suggesting that participants were not systematically biased by the temporal structure of baiting (i.e., they did not systematically prefer the first or last baited cup). Moreover, success in Experiment 1 aligns closely with prior studies using different occlusion or baiting procedures, including paradigms in which both cups were hidden behind a screen, further suggesting that these effects are not driven by superficial features of the hiding procedure.

An increased desire to explore and seek information relative to other primates—even at the cost of a reward—has been proposed as an additional reason why children may exhibit lower-than-ceiling performance on modal tasks (Cesana-Arlotti et al., 2025; Lapidow & Walker, 2020; Sanchez-Amaro & Rossano, 2023). A prior study indicated that this may not be the case through deploying a search time paradigm (Andreuccioli et al., 2024), and we believe that Experiment 3 offers another counter to this proposition; as the guaranteed item here was a non-reward, the participant did not stand to reduce any uncertainty from choosing this option. Nevertheless, a quarter of Experiment 3's subjects did so, indicating that failures in these tasks may not be attributable to curiosity.

Overall, the patterns of performance found in Experiments 1 and 2 cannot be easily distinguished from levels of success predicted by the minimal representation account, which predicts that three-year-old children generally only simulate one possible outcome at a time and treat it as reality until proven otherwise (Leahy & Carey, 2020). One open question involves how a minimal representer would handle a challenge such as Experiment 3 if the subject simulates the lone reward being removed; future work may benefit from more systematically investigating how children make decisions when confronted with only impossible reward outcomes. Additional investigations into the memory limits of three-year-olds on structurally similar tasks not containing possibilities could further elucidate what true ceiling performance looks like—macaques' success rate of about 80% in successful experiments was not equivalent to 100%, but it was precisely on par with a purely memory-based task (Braswell, Meade, et al., 2025).

In sum, while our findings suggest that early successes in modal reasoning may depend less on access to modal language than on the structure of the possibilities represented within a task; however, the strict constraints observed in rhesus macaques may be less deterministic for early human cognition, instead giving way to more flexible—but nonetheless limited—forms of reasoning about possibility.

Acknowledgements

We thank Alexis Burke and Rita Usmanova for assistance with data collection. Additionally, we thank the members of the Comparative Cognition Lab for their advice and feedback on the studies. We are grateful to Birch Aquarium at Scripps and Fleet Science Center for their longstanding support in facilitating data collection. Finally, we thank our participants and families for taking part in our studies.

References

Alderete, S., & Xu, F. (2023). Three-year-old children's reasoning about possibilities. *Cognition*, 237, 105472. https://doi.org/10.1016/j.cognition.2023.105472

Andreuccioli, L., Mazor, S., Begus, K., Bonawitz, E., Denison, S., & Walker, C. M. (2024). Young children adapt their search behavior for necessary versus merely possible outcomes. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 46.

Braswell, N., Meade, M. G., Cesana-Arlotti, N., & Santos, L. (2025). Wild possibilities: Evidence of modal cognition in free-ranging rhesus macaques (Macaca mulatta). In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 47)

Brody, G., Mazalik, P., & Feiman, R. (2024). Object files encode possible object identities, but not possible locations. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 46).

Cesana-Arlotti, N., & Halberda, J. (2024). A continuity in logical development: Domain-general disjunctive inference by toddlers. *Open Mind*, 8, 809–825.

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), e70044.

Cesana-Arlotti, N., Kovács, Á. M., & Téglás, E. (2020). Infants recruit logic to learn about the social world. *Nature Communications*, 11(1), 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.

Engelmann, J. M., Haux, L. M., Völter, C., Schleihauf, H., Call, J., Rakoczy, H., & Herrmann, E. (2023). Do chimpanzees reason logically? *Child Development*, 94(5), 1102–1116. https://doi.org/10.1111/cdev.13861

Engelmann, J. M., Völter, C. J., Goddu, M. K., Call, J., Rakoczy, H., & Herrmann, E. (2023). Chimpanzees prepare for alternative possible outcomes. *Biology Letters*, 19(6), 20230179.

Engelmann, J. M., Völter, C. J., O'Madagain, C., Proft, M., Haun, D. B. M., Rakoczy, H., & Herrmann, E. (2021). Chimpanzees consider alternative possibilities. *Current Biology*, 31(20), R1377–R1378. https://doi.org/10.1016/j.cub.2021.09.012

Ferrigno, S., Huang, Y., & Cantlon, J. F. (2021). Reasoning through the disjunctive syllogism in monkeys. *Psychological Science*, 32(2), 292-300.

Gautam, S., Suddendorf, T., & Redshaw, J. (2021). Do monkeys and young children understand exclusive "or" relations? A commentary on Ferrigno et al.(2021). *Psychological Science*, 32(11), 1865-1867.

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.

Johnson-Laird, P. N., Byrne, R. M., & Khemlani, S. S. (2024). Models of possibilities instead of logic as the basis of human reasoning. *Minds and Machines*, 34(3), 19.

Lapidow, E., & Walker, C. M. (2020). Informative experimentation in intuitive science: Children select and learn from their own causal interventions. *Cognition*, 201, 104315.

Leahy, B. (2024). Many preschoolers do not distinguish the possible from the impossible in a marble-catching task. *Journal of Experimental Child Psychology*, 238, 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*, 119(52), e2207499119.

Leahy, B. P., & Carey, S. E. (2020). The acquisition of modal concepts. *Trends in Cognitive Sciences*, 24(1), 65–78.

Leahy, B., & Zalnieriunas, E. (2021). Might and might not: Children's conceptual development and the acquisition of modal verbs. *Semantics and Linguistic Theory*, 426–445.

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

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

Redshaw, J., & Suddendorf, T. (2024). Can chimpanzees conceive of mutually exclusive future possibilities? A Comment on:'Chimpanzees prepare for alternative possible outcomes'(2023), by Engelmann et al. *Biology Letters*, 20(6), 20230409.

Roese, N. J., & Olson, J. M. (2014). Counterfactual thinking: A critical overview. What Might Have Been, 1–55.

Samet, D. (1996). Hypothetical knowledge and games with perfect information. *Games and Economic Behavior*, 17(2), 230–251.

Sánchez-Amaro, A., & Rossano, F. (2023). Comparative curiosity: How do great apes and children deal with uncertainty? *PLoS One*, 18(5), e0285946.

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, 106078.

Van Hoeck, N., Begtas, E., Steen, J., Kestemont, J., Vandekerckhove, M., & Van Overwalle, F. (2014). False belief and counterfactual reasoning in a social environment. *NeuroImage*, 90, 315–325