# The origin of the possible: 12-month-olds' understanding of certain, likely, and unlikely events

**Şeref Can Esmer (seref.esmer@yale.edu)**
Department of Psychology, 100 College Street,
New Haven, CT 06510 USA

**Nicolò Cesana-Arlotti (nicolo.cesana-arlotti@yale.edu)**
Department of Psychology, 100 College Street,
New Haven, CT 06510 USA

**Abstract**

To predict and prepare for near-future outcomes, infants must respond to the variability in their probability. Adults achieve that with modal concepts that quantify over multiple possibilities, but whether and how infants can do the same is unclear. In two preregistered habituation experiments, we asked whether infants can distinguish outcomes based on physical probability level (100% vs. 66% in Experiment 1. 66% vs. 33% in Experiment 2). 12-month-olds were habituated to events with 66% probability, and their proportion of looking at 100% (Exp 1., N=35) and 33% (Exp 2., N=24) events were measured before (i.e., baseline) and after habituation (i.e., test). We found that infants' proportion of looking at events with 33% probability (in Exp 2), but not at events with 100% probability (in Exp 1), increased from baseline to test. Thus, 12-month-olds distinguish likely events from unlikely ones but not from necessary events.

**Keywords:** representation of possibilities, infancy, logic, probabilistic reasoning, cognitive development

## Introduction

To understand and learn about the world surrounding them, babies form expectations about what may happen next (Téglás et al., 2007; Schulz, 2012; Stahl & Feigenson, 2015). Prediction is challenging because not all future outcomes are equiprobable. While some events have to happen (e.g., sunset), some others may or may not happen (e.g., rain). Further, events can also differ based on their probability of happening – their modal status (Hacquard, 2011). Some events are likely to happen (more likely than not to happen, like rain in monsoon season), while others are unlikely to happen (less likely than not to happen, like rain in the desert). Thus, possible events in our environment vary vastly regarding their probability. Such vast differences could make prediction and learning difficult. However, developmental psychologists have learned that children, and even infants, seem to efficiently calibrate their exploration (Schulz, 2012; Schulz & Bonawitz, 2007), and expectations (Xu & Garcia, 2008; Téglás et al., 2007) to the probability of events in the environment. These results raise profound questions. How does a baby without any training in formal probability theory or relevant linguistic knowledge (Leahy & Zalnieriunas, 2022; Redshaw & Ganea, 2022; Leahy & Carey, 2020; Ozturk & Papafragou, 2015) think about varying probability of events? The present study aims to shed light on the foundation of modal cognition: the representations available to infants to distinguish events and future possibilities based on their modal status.

To respond to the variability of events' probability, adults can categorize them with modal concepts. Classical modal concepts categorize possible events via classic quantification over alternative possibilities (Kripke, 1963). Thus, an event is necessary if it is realized in all possible outcomes, merely possible if manifest in only some, or impossible if manifested in none. Alternatively, outcomes can be categorized with comparative modal concepts via comparative quantification of possibilities (Johnson-Laird et al., 1999; Cesana-Arlotti et al., 2012). For example, a future event is likely if it is manifested in most of the possible outcomes or unlikely if it fails to happen in most of them. In this study, we start to explore which modal concepts, if any, shape infants' expectations about the uncertain future.

Infants show differential reactions to improbable and probable events. For example, when presented with a lottery box composed of three yellow and one blue object, 12-month-olds looked longer at the outcome when it was the blue object (i.e., improbable) than when it was one of the yellow ones (i.e., probable; Téglás et al., 2007). Similarly, even 6- and 8-month-olds looked longer at samples that were less likely to come from an observed population (Denison et al., 2013; Xu & Garcia, 2008). Further, infants also explored boxes from which unlikely samples were drawn (Sim & Xu, 2017). These probabilistic intuitions can interact with other physical cues (e.g., proximity to the exit of the box, Téglás et al., 2011; immobile objects that cannot be sampled, Denison et al., 2014) and psychological cues (e.g., whether the agent is intentionally or randomly sampling, Xu & Denison, 2009; whether the experimenter was surprised at an unobserved outcome, Wu et al., 2024). All these findings motivate the proposal that infants have the cognitive resources to distinguish the probabilities of events, integrate subtle cues, and look longer at improbable events than probable ones.

Leahy and Carey (2020) argued that infants' sensitivity to an outcome probability doesn't necessarily require the representation of modal concepts (or probabilities). Instead, infants' responses might reflect a minimal representation of possibility. According to this proposal, children can only simulate a random single possible outcome when predicting an event and they cannot represent multiple possibilities before they acquire modal words like "maybe" or "or". This strategy was used to capture previous findings like those of Mody and Carey (2016) or Téglás et al. (2007), among others. For example, when presented with a physical lottery, like the



In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* 

one of Téglás et al. (2007), infants would simulate a single possible outcome; more often a probable one than an improbable one, resulting in higher group-level average looking time at the latter (Leahy & Carey, 2020).

More broadly, if infants and young children can only think of just a single possibility at the time, as Leahy and Carey (2020) argued, and cannot even represent a small set of possible outcomes, they shouldn't be able to use modal concepts. Indeed, concepts like "guaranteed," "likely," "unlikely," or "impossible" require the representation of multiple, alternative possibilities to judge whether an event takes place in all, most, least, or none of them.

However, new findings invite a reappraisal of the minimalistic account of infant and young children's probabilistic intuitions. First, young preschoolers may be capable of deploying classic modal concepts to evaluate the outcome of a probabilistic event. For example, Cesana-Arlotti et al. (under review) showed that more 3-year-olds chose a machine with 100% probability than 66% probability when the outcome was wanted and refrained from it when the outcome was unwanted. However, this pattern was not observed when the machines had 33% and 66% probabilities. Hence, 3-year-olds distinguished certain options from merely possible ones but didn't differentiate between two merely possible options with different probabilities. Further, 2- and 3-year-olds learned better from impossible events than very unlikely but still possible events (Stahl & Feigenson, 2024). Thus, young children may deploy classic modal concepts in their decisions and learning.

Second, a pupillometric study with 14-month-old infants found that infants can represent multiple possibilities (Cesana-Arlotti et al., 2022). In the study, 14-month-olds showed higher pupil dilation, an index of cognitive effort, to uncertain events (i.e., the identity of the hidden object is compatible with two possibilities) than certain events (i.e., the identity of the hidden object is compatible with only one possibility; Cesana-Arlotti et al., 2022). If, in that experiment, infants can think of only a single outcome, regardless of the open possibilities, the cognitive effort should have been similar.

Therefore, recent studies suggested that preschoolers may deploy modal concepts before mastering words like "can" and "has to" (Leahy & Zalnieriunas, 2022), and infants may represent multiple possibilities, raising an important question: can infants also use modal concepts? If so, what kind of modal concepts are available to them? To address this question, we ran two preregistered experiments with 12-month-olds. In the experiments, infants were presented with multiple lottery events (as in Téglás et al, 2007, 2011, 2015), but in three phases. In the baseline phase, we presented events with different modal status (i.e., Necessary-100% vs. Possible-66% in Experiment 1; Likely-66% vs. Unlikely-33% in Experiment 2) to measure if infants had an initial looking preference for a particular event type (as in Göksun et al., 2011; Pruden et al., 2013). Then, infants were habituated to events with 66% probability (Possible-66% in Experiment 1 and Likely-66% in Experiment 2) that varied in the types of objects (i.e., colors and shapes). Finally, in the test phase, infants were again shown the same events used in the baseline phase to measure the impact of habituation on their looking preference. To account for potential initial biases for a specific category or probability level, we used the change (i.e., difference) in looking at proportion from baseline to test as a measure of sensitivity to modal contrasts. (i.e., Necessary-100% vs. Possible-66% in Experiment 1; Likely-66% vs. Unlikely-33% in Experiment 2). If infants can distinguish possible events based on their relative modal status, they would look relatively longer at the nonhabituated event types (Necessary-100% in Experiment 1 and Unlikely-33% in Experiment 2) at the test phase than at the baseline.

## Experiment 1

In Experiment 1, we asked whether infants' proportion of looking at Necessary-100% events increased after habituation to Possible-66% events.

### Method

Both experiments' sample size criteria, exclusion criteria, analyses, and key experimental parameters were preregistered. The preregistrations are publicly available at: https://osf.io/4haz7/?view_only=c1f35d5562624eecac0aafd97d0a7cf0.

**Participants.** A total of 35 full-term healthy infants were included in the analyses ($M_{age}$ = 11m 29d, *SD* = 17d, Range = 11m 5d - 13m 1d; 16 girls). Additional infants were tested but excluded due to equipment failure (N = 1), lack of enough valid trials (N = 7), or experimenter error (N = 1). The study was approved by the Institutional Review Board of Yale University.

We used a Sequential Bayes Factor procedure to decide the final sample size (Mani et al., 2021). Our starting sample size is 24 11- to 13-month-olds. Once we reached this sample size, we ran a Bayesian paired-sample t-test to compare the looking proportion at the Necessary-100% events at the baseline and the test (see results). If the Bayes Factor (BF) was between 1/3 and 5, we continued data collection and ran the same analysis after each new participant. We stopped data collection if we reached BF < 1/3 or BF > 5, or a sample of 48 participants.

**Stimuli**. In all the movies presented, infants saw a box with objects bouncing inside (see Fig. 1). The box had a horizontal divider inside that divided it into two chambers. Only the lower chamber was spatially connected to an exit (a hole) at its bottom. One object was placed above the divider so it could not exit the box. After bouncing for 10 s, the curtains close to occlude the content of the box. Two seconds later, one of the objects exited the box, moving for 0.75s, and stopped outside the box for the rest of the trial.

Each infant watched two familiarization movies, followed by four baseline movies, then 6 to 14 habituation trials, and then up to 8 test movies.

In the familiarization movies, infants saw a lottery box where a blue cube, a red ball, and a pink pyramid were bouncing. The pink pyramid was placed above the divider and the others below it. Infants watched the exit of the red ball in one familiarization trial and the blue cube in the other. The order of presentation was alternated across participants.

In the four baseline movies, infants saw either a lottery box with three blue cubes and a red ball or one with three red balls and a blue cube. Baseline trials differed based on the modal status of the outcome (Necessary-100% vs. Possible-66%). In Necessary-100% events, the singleton object (e.g., one blue cube) was placed above the divider, whereas the three objects in the chamber connected to the exit were identical (e.g., three red balls), making the outcome of a red ball guaranteed. In the Possible-66% events, one of the three identical objects (e.g., one red ball) was placed above the divider, whereas in the chamber connected to the exit, there were two identical objects, and the singleton one (e.g., two red balls and one blue cube), making the outcome of a red ball probable but not guaranteed (66% chance).

The order of presentation of Necessary-100% and Possible-66% events followed an ABBA sequence, with half of the participants starting with Necessary-100%. In the first and second half of the baseline phase, the majority object type bouncing inside the box and, therefore, the outcome objects were different. Half of the participants saw blue cubes as the majority object type and outcome in the first two trials, and the remaining saw red balls.

In the habituation trials, four pairs of object types (i.e., light blue cylinder with claret red holed cube, purple donut with yellow diamond, light pink plus sign with dark gray icosphere, and green cone with pink pyramid) bounced inside the box. The distribution of the objects inside the box and the outcome followed the Possible-66% event type. In every four habituation trials, each pair of object types was presented exactly one time. With every repetition of the same pair, the role of the object types in a pair was switched (i.e., the type of singleton object and the type of the triplet objects). We prepared four orders of presentation of the object pairs for the habituation (i.e., P1-P2-P3-P4, P2-P3-P4-P1, P3-P4-P1-P2, P4-P1-P2-P3) that were counterbalanced across participants.

In the test trials, the same Necessary-100% and Possible-66% events used in the baseline trials were presented in an ABABABAB order. Half of the infants started watching Necessary-100% events, while the remaining started with Possible-66% events. Infants who saw blue cubes (or red balls) exiting in the first two baseline trials saw blue cubes (or red balls) as the outcomes in the first two test trials.

Before each movie, an attention-getter (colorful arrows pointing to the center of the screen with a vibrant sound) was presented to ensure infants attended the movie from the start.

Thirty-two experimental conditions were created by crossing two baseline orders (NPPN or PNNP), four habituation orders, and two test orders (NPNPNPNP or PNPNPNPN).

**Procedure**. The experiment was conducted online over a Zoom video call. Infants' caregivers were sent the Slides.com link that involves the stimuli, and infants saw the stimuli in their home environment from the caregivers' computers. Caregivers shared their screens with the experimenter so that the experimenter could monitor the stimuli presentation. Zoom video panels and floating meeting controls were hidden from the infants' view. Infants sat on their caregivers' lap at a distance from which they could not reach the computer. The caregivers were instructed to close their eyes and not interfere with the experiment procedure. The procedure was screen-recorded for offline coding.

The stimuli presentation was controlled by using the PyHab online edition (Kominsky et al., 2021). To ease the online and offline coding, infants were first presented with a 5-point calibration procedure (i.e., center and four corners). Then, they proceeded to the experiment blocks (i.e., familiarization, baseline, habituation, and test). Infants' looking times were online coded in the baseline, habituation, and test trials. Looking coding started when the outcome object started to appear in the tube. The trial ended if the child looked away for two consecutive seconds or 30s after the outcome object was first revealed inside the tube. Trials were repeated if infants did not look at the bouncing objects before the curtains closed for at least 1s and if the trials were invalid according to preregistered exclusion criteria (see below).

The number of habituation trials ranged from 6 to 14 depending on the infants' looking times. The habituation procedure ended if the average looking time at the last three habituation trials dropped below half of the average looking times of the first three habituation trials.

Infants' looking times were coded offline using DataVyu and then used in the analyses. Baseline and test trials were excluded if (i) the caregiver interferes verbally or behaviorally, (ii) the experimenter ends the trial before the baby looks away for 2 s, (iii) looking times are below 1 s or above/below 2.5 absolute median deviations around the median by experimental condition, (iv) the child gets fussy, and (v) equipment malfunction interferes with the experimental procedure.

## Results and Discussion

Experiment 1 asked if infants' proportion of looking at Necessary-100% events increased after habituation to Possible-66% events. We started by averaging each participant's looking time by event type (Necessary-100% vs. Possible-66%) and experiment phase (baseline vs. test). Based on the preregistration, we included only the first four test trials. Next, for each participant and experiment phase, we computed a *proportion of looking* at the Necessary-100% event by dividing the mean looking time at the Necessary-100% events by the sum of the mean looking time at the Necessary-100% events and the Possible-66% one.

Based on the preregistered stopping criteria, we achieved BF10 < 1/5 with a sample of $N = 35$ participants. Shapiro-Wilk tests on the distribution of the proportion of looking at Necessary-100% events at baseline did not detect a deviation from normality ($W = .976$, $p = .625$) but detected it at test ($W = .929$, $p = .026$). However, the difference in the proportion of looking at Necessary-100% events from baseline to test did not deviate from normality ($W = .959$, $p = .216$).

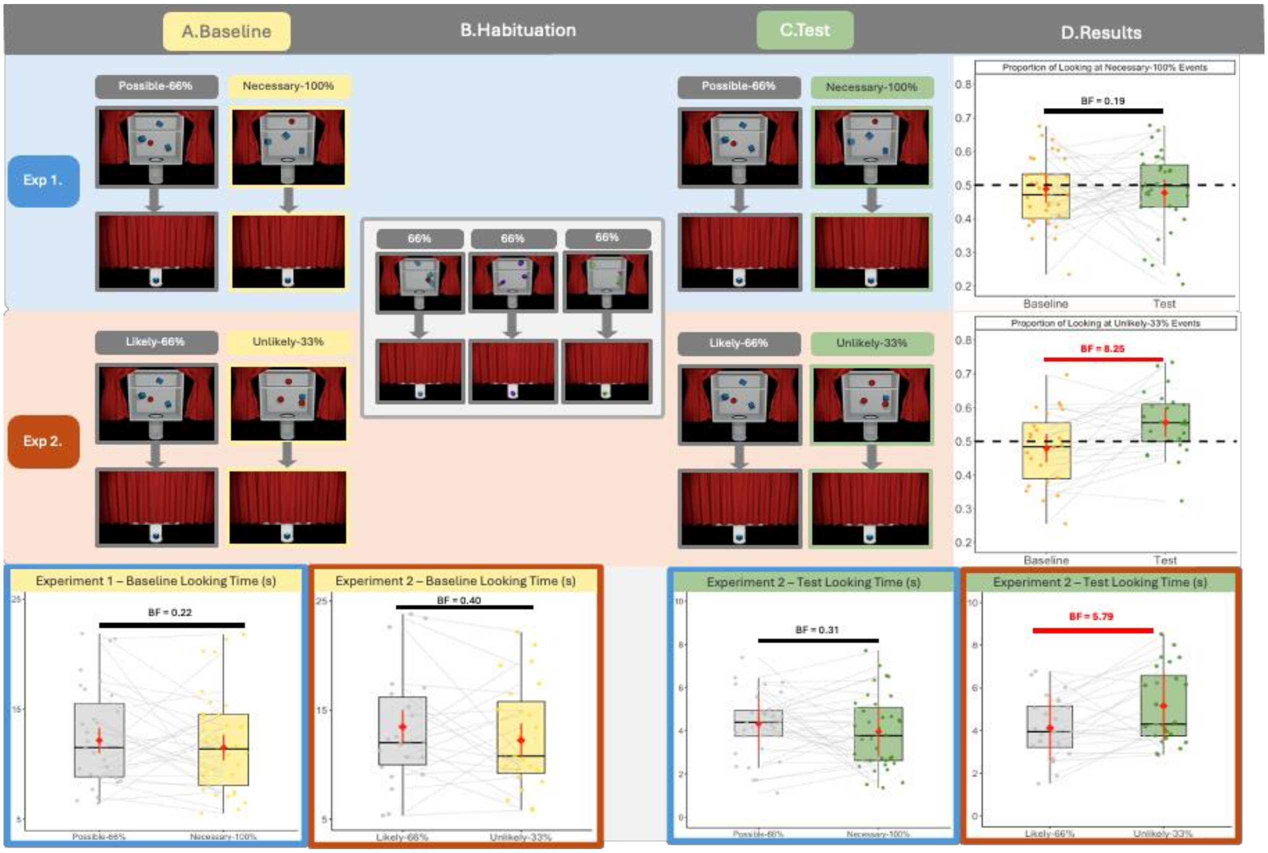


Figure 1: The structure, example stimuli, and results of Experiment 1 (upper panel, gray) and Experiment 2 (middle panel, orange). Panel A shows the example stimuli used in the baseline phase of the experiments (Possible vs. Necessary in Experiment 1, Likely vs. Unlikely in Experiment 2). Panel B shows three different example stimuli from the habituation phase. Panel C shows example test stimuli, which are identical to the baseline examples. Panel D shows the proportion of looking at the nonhabituated event (Necessary-100% in Exp 1, Unlikely-33% in Exp 2) in the baseline and test phases. The bottom panel shows the looking times (in seconds) at each event type at the baseline (left panel) and the test phase (right panel). Error bars represent 95% confidence intervals; BF stands for Bayes Factor. BF > 3 indicates moderate evidence that measures are different from each other and are presented in bold and red. BF<0.33 indicates moderate evidence that measures are comparable.

Thus, we ran a Bayesian t-test comparing looking proportions at baseline and test phases using the t-testBF function from the BayesFactor package (Morey et al., 2015). The test found moderate evidence of no difference between looking proportions at Necessary-100% events at baseline and test phases ($M_{\text{Baseline}}$ = 0.49, $SD$ = 0.11; $M_{\text{Test}}$= 0.48, $SD$ = 0.13; BF = 0.19). A t-test didn't detect a difference between the two looking proportions ($t(34)$ = -.036, $p$ = .723). This result suggests that infants didn't distinguish Possible-66% events from Necessary-100% events.

Next, we ran exploratory analyses to further characterize infants' looking times. First, we asked if infants' looking times at Necessary-100% and Possible-66% differed in the baseline phase to check if they had an initial preference for a type of event. Shapiro-Wilk tests on the distribution of looking times at the baseline trials did not detect deviations from normality (Necessary-100%-$W$ = .945, $p$ = .079; Possible-66%- $W$ = .944, $p$ = .075). Thus, we ran a Bayesian paired t-test to compare looking times at Necessary-100% and Possible-66% in the baseline phase. The test found moderate evidence that looking times were not different in the two conditions ($M_{\text{Necessary-100\%}}$ = 11.5s, $SD$ = 4.7s; $M_{\text{Possible-66\%}}$= 12.2s, $SD$ = 5.3s, BF=0.22). A t-test didn't detect a difference between the looking times at the two types of events in the baseline ($t(34)$ = -.679, $p$ = .501).

Last, we asked whether infants' looking times at Necessary-100% and Possible-66% differed in the test phase. Shapiro-Wilk tests on the distribution of looking times at the test trials did not detect deviations from normality (Necessary-100%-$W$ = .959, $p$ = .210; Possible-66%- $W$ = .975, $p$ = .596). Thus, we ran a Bayesian paired t-test to compare looking times at Necessary-100% and Possible-66% in the test phase. The test detected moderate evidence that looking times were not different in the two conditions ($M_{\text{Necessary-100\%}}$ = 4.0s, $SD$ = 1.7s; $M_{\text{Possible-66\%}}$= 4.3s, $SD$ = 1.5s, BF = 0.31). A t-test didn't detect a difference between the looking times at the two types of events in the test phase ($t(34)$ = -1.094, $p$ = .281). This result confirms that, after habituation to the Possible-66% events, infants do not change their looking time at Necessary-100%, suggesting a lack of discrimination between the Possible-66% and Necessary-100%.

## Experiment 2

In Experiment 2, we asked whether infants' proportion of looking at Unlikely-33% events increased after habituation to Likely-66% events.

### Method

**Participants** A total of 24 full-term healthy infants were included in the analyses ($M_{age}$ = 11m26d, *SD*=15d Range =11m2d-12m20d; 10 girls). Additional infants were tested but excluded due to equipment failure (N = 2), for not providing enough valid trials (N = 3) or being preterm born (N=1). Similar to Experiment 1, a Sequential Bayes Factor procedure was used to determine sample size.

**Stimuli and Procedure** Stimuli and procedure were identical to Experiment 1 except for the baseline and test phases. In the baseline and test phase, Likely-66% and Unlikely-33% events were presented. Likely-66% events were identical to Possible-66% events in Experiment 1. The distribution of the objects inside the lottery box in Unlikely-33% events is identical to the Likely-66%, but the outcome object is from the minority category.

### Results and Discussion

Experiment 2 asked if infants' proportion of looking at Unlikely-33% events increased after habituation to Likely-66% events. We started by averaging each participant's looking time by event type (Unlikely-33% vs. Likely-66%) and experiment phase (baseline vs. test). Based on the preregistration, we only used the first four test trials to calculate the measure for the test phase. Next, for each participant and experiment phase, we computed a *proportion of looking* at the Unlikely-33% event by dividing the mean looking time at the Unlikely-33% event by the sum of the mean looking time at the Unlikely-33% event and the Likely-66% one.

Based on the pre-registered stopping criteria, we achieved $BF_{10} > 5$ with a sample of $N$ = 24 participants. Shapiro-Wilk tests on the distribution of the proportion of looking at Unlikely-33% events at baseline, and at test did not detect a deviation from normality (Baseline-$W$ = .98, $p$ = .949; Test-$W$ = .982, $p$ = .931). Thus, we ran a Bayesian t-test to compare looking proportions at the baseline and test phases. The test found moderate evidence of a difference between looking proportions at Unlikely-33% events, which increased from baseline to test ($M_{\text{Baseline}}$ = 0.48, $SD$ = 0.11; $M_{\text{Test}}$= 0.56, $SD$ = 0.09; BF = 8.25). A t-test also detected a difference between the two looking proportions ($t$(23) = -3.077, $p$ = .005). This result suggests that infants were habituated to the Likely-66% events and, as a result, their proportion of looking at the Unlikely-33% events increased from baseline to test.

Next, we ran exploratory analyses to further characterize infants' looking times. First, we asked if infants' looking times at Unlikely-33% and Likely-66% differed in the baseline phase to check if they had an initial preference for a type of event. Shapiro-Wilk tests on the distribution of looking times at the baseline trials did not detect deviations from normality (Unlikely-33%-$W$ = .96, $p$ = .405; Likely-66%- $W$ = .930, $p$ = .098). Thus, we ran a Bayesian paired t-test to compare looking times at Unlikely-33% and Likely-66% in the baseline phase. The test found anecdotal evidence that looking times were not different in the two conditions ($M_{\text{Unlikely-33\%}}$ = 12.2s, $SD$ = 5.1s; $M_{\text{Likely-66\%}}$= 13.5s, $SD$ = 6.1s, BF=0.40). A t-test didn't detect a difference between the looking times at the two types of events in the baseline ($t$(23) = -1.179, $p$ = .251).

Last, we asked whether infants' looking times at Unlikely-33% and Likely-66% differed in the test phase. Shapiro-Wilk tests on the distribution of the difference in looking times at unlikely and likely events at the test trials did not detect deviations from normality ($W$ = .98, $p$ = .851). Thus, we ran a Bayesian paired t-test to compare looking times at Unlikely-33% and Likely-66% in the test phase. The test detected moderate evidence that infants looked longer at the Unlikely-33% events than Likely-66% ($M_{\text{Unlikely-33\%}}$ = 5.2s, $SD$ = 1.8s; $M_{\text{Likely-66\%}}$ = 4.1s, $SD$ = 1.4s, BF = 5.79) events at the test phase. A t-test also detected a difference between the looking times at the two types of events in the baseline ($t$(23) = 2.899, $p$ = .008). This result confirms that, after habituation to the Likely-66% events, infants look longer at new instances of the Unlikely-33% event than those of the Likely-66% event. Thus, infants could distinguish the events with 66% probability from the ones with 33% probability.

## Discussion

Our study asked whether 12-month-old infants can distinguish different events based on representations of modal status. Infants were presented with Necessary-100% and Possible-66% events in Experiment 1, and Likely-66% and Unlikely-33% events in Experiment 2. They watched these events before (i.e., baseline) and after (i.e., test) habituation to events with 66% probability. If infants can represent the different modal status of two types of events (e.g., Likely-66% and Unlikely-33%), the proportion of looking to the one event type (e.g., Unlikely-33%) should increase after habituation to the other type (e.g., Likely-66%). Our results revealed a striking pattern: infants distinguished events with 66% probabilities from 33% probabilities in Experiment 2 but not from 100% probabilities in Experiment 1.

This pattern suggests infants have the representational resources to distinguish likely from unlikely events. First, the lack of differentiation between necessary and possible events in Experiment 1 speaks against the possibility that infants simply responded to low-level properties in Experiment 2 (e.g., the number, color, and distribution of the objects inside the container or the number of objects of the same type as the outcome).

Moreover, our results also speak against the hypothesis that infants used the minimal representation of possibility in our task (Leahy & Carey, 2020). If infants had just simulated a single possibility in the test trials, as the minimal account predicted, we should have observed a similar pattern in Experiments 1 and 2. According to this account, all participants would have simulated the correct outcome in Necessary-100% events, 66% of participants simulated the correct outcome in Likely-66%, and 33% of participants simulated the correct outcome in Unlikely-33% events. Thus, the difference in the proportion of participants surprised by the outcomes should have been the same in Experiment 1 and 2. Thus, it is unclear how the minimal representation of possibility could account for the increase in looking proportion at Likely-66% events we have recorded only in Experiment 2 but not in Experiment 1.

Furthermore, our results sharply contrast with the preschoolers' performance in a recent modal decision-making experiment (Cesana-Arlotti et al., under review). In this study, children succeeded in choosing between options (probabilistic machines) with probabilities that crossed classical modal categories (i.e., impossible, merely possible, necessary) but failed between equidistant probabilities within the same classic category (e.g., 33% vs. 66%, both mere possibilities). Remarkably, this pattern is exactly the opposite of what we observed in 12-month-old infants in our Experiments: they differentiated two merely possible events crossing categories unlikely/likely (33% vs. 66%) but failed to distinguish possible from necessary events when they were both likely outcomes (66% vs. 100%). So, our results align with the intriguing hypothesis that infants and preschoolers may prioritize different modal concepts.

Therefore, a critical question emerges from our findings: what kind of modal concepts may carve the logical space of possibilities in the mind of a one-year-old baby?

A first possibility is that infants represent the probability of an observed event as a ratio between the possible outcomes realizing it and the full set of possible outcomes (under an assumption of equiprobability, Johnson-Laird et al., 1999); and thus the discriminability of two ratios determine the discriminability of probabilistic events. Crucially, infants' successful differentiation in Experiment 2 and lack of it in Experiment 1 can be explained if infants compare events modal status based on probability ratio comparison. Indeed, Experiment 1 has a less favorable difference in probability ratio (1/0.66=1.5) than Experiment 2 (.66/.33=2, i.e., the higher the ratio, the more different the events), making it more difficult to discriminate the events in Experiment 1 than Experiment 2. In line with this account, infants can discriminate between two ratios when they are as distant as the probabilities in Experiment 2 but not when as close as the ones in Experiment 1 (McCrink & Wynn, 2004). Interestingly, similar sensitivity to ratio changes has also been observed in chimpanzees' probabilistic judgments. In a study by Hanus & Call, 2014, chimpanzees were more likely to choose trays with higher probability rewards when the probability ratio changes (high probability/low probability) were more marked, particularly two or higher.

Alternatively, rather than computing probabilities, infants might have learned the heuristic that the outcome object was always the same kind as *the most common object kind* inside the box (e.g., if there are more red balls than blue cubes, then it will be a red ball). This explanation is compatible with our data such that the outcomes of both Necessary-100% and Possible-66% events were from the most common object kind, while the outcomes of the Unlikely-33% events were not. As a result, infants might have preferred to look longer at Unlikely-33% events, where the learned heuristic did not work. However, previous studies using similar events found that infants expected the most common object to exit from the box only when it was the most probable outcome (i.e., when the duration of the occlusion was sufficient to make the last seen proximity to the exit irrelevant; Téglás et al., 2011), suggesting that infants are not simply responding to a heuristic.

Furthermore, a third possibility is that the infants may deploy modal concepts of the comparative kind: an event could be categorized as *likely* (more likely to happen than not) or *unlikely* (less likely to happen than not). Crucially, these concepts predict the infants' response observed in our experiments. Both necessary and merely possible but likely events belong to *the likely* category, while possibilities with a probability lower than 50% belong to the distinct category of *the unlikely*. Thus, this novel *unlikely* event may have elicited more attention in the test phase than the habituated *likely* events (i.e., likely). Given these different competing explanations, future experiments are required to adjudicate the nature of infants' representations of likely and unlikely events.

Lastly, our experiments recorded a second pattern that differed from previous studies. Infants in our study looked at events that differed in probabilities in the baseline phase equally. In contrast, previous studies found that infants look longer at improbable sampling events than at probable ones (Denison et al., 2013; Téglás et al., 2007; Téglás et al., 2011; Xu & Garcia, 2008). In these studies, infants saw sampling outcomes together with the sampling population. Unlike previous studies, in our study, the population (i.e., lottery box) was occluded while viewing the outcome object, which may have made it difficult to compare the outcome and the population distribution in our study.

In conclusion, our results recorded a striking pattern: 12-month-olds distinguished 66%-probability events from 33%-probability ones but lumped them together with 100%-probability events, as indexed by the change in their proportion of looking at particular events. This pattern of results is unexpected if infants can only simulate a single possible outcome at the time (Leahy & Carey, 2020). The results are also unexpected if infants use classical modal categories like *the necessary* and *the possible*. Thus, our results open the questions of the nature of the mental representations that establish the distinction between *the likely* and *the unlikely* in the human infant mind.

## Acknowledgments

We would like to thank all families who participated in the study, as well as Mahham Fayyaz, Nikita Shtarkman, Griffen

Malkin, Molly Atiencia, and Clariss Bolanos for help in stimulus preparation, participant recruitment, and data coding.

## References


Cesana-Arlotti, N., Jáuregui, S., Mazalik, P., Nichols, S., & Halberda, J. (2025). *Logical concepts of (im)possibility guide young children's decision-making.* Manuscript submitted for publication.

Cesana-Arlotti, N., Téglás, E., & Bonatti, L. L. (2012). The probable and the possible at 12 months: Intuitive reasoning about the uncertain future. *Advances in Child Development and Behavior*, *43*, 1-25.

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 B*, *377*(1866), 20210343.

Denison, S., Reed, C., & Xu, F. (2013). The emergence of probabilistic reasoning in very young infants: evidence from 4.5-and 6-month-olds. *Developmental Psychology*, *49*(2), 243-249.

Denison, S., Trikutam, P., & Xu, F. (2014). Probability versus representativeness in infancy: Can infants use naïve physics to adjust population base rates in probabilistic inference?. *Developmental Psychology*, *50*(8), 2009-2019.

Göksun, T., Hirsh-Pasek, K., Golinkoff, R. M., Imai, M., Konishi, H., & Okada, H. (2011). Who is crossing where? Infants' discrimination of figures and grounds in events. *Cognition, 121*(2), 176-195.

Hacquard, V. (2011). Modality. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.). *Semantics: An international handbook of natural language meaning* (pp. 1484–1515). Berlin: Mouton de Gruyter

Hanus, D., & Call, J. (2014). When maths trumps logic: probabilistic judgements in chimpanzees. *Biology Letters*, *10*(12), 20140892.

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M. S., & Caverni, J. P. (1999). Naive probability: a mental model theory of extensional reasoning. *Psychological Review*, *106*(1), 62.

Kominsky, J. F., Begus, K., Bass, I., Colantonio, J., Leonard, J. A., Mackey, A. P., & Bonawitz, E. (2021). Organizing the methodological toolbox: Lessons learned from implementing developmental methods online. *Frontiers in Psychology*, *12*, 702710.

Kripke, S. A. (1963). Semantical analysis of modal logic i normal modal propositional calculi. *Mathematical Logic Quarterly, 9*(5-6), 67-96.

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. In *Semantics and Linguistic Theory* (pp. 426-445).

Mani, N., Schreiner, M. S., Brase, J., Köhler, K., Strassen, K., Postin, D., & Schultze, T. (2021). Sequential Bayes Factor designs in developmental research: Studies on early word learning. *Developmental Science*, *24*(4), e13097.

McCrink, K., & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychological Science*, *18*(8), 740-745.

Mody, S., & Carey, S. (2016). The emergence of reasoning by the disjunctive syllogism in early childhood. *Cognition*, *154*, 40-48.

Morey, Richard D., Jeffrey N. Rouder, Tahira Jamil, and Maintainer Richard D. Morey. Package '*bayesfactor*'. 2015.

Ozturk, O., & Papafragou, A. (2015). The acquisition of epistemic modality: From semantic meaning to pragmatic interpretation. *Language Learning and Development*, *11*(3), 191-214.

Pruden, S. M., Roseberry, S., Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Infant categorization of path relations during dynamic events. *Child Development, 84*(1), 331-345.

Redshaw, J., & Ganea, P. A. (2022). Thinking about possibilities: mechanisms, ontogeny, functions and phylogeny. *Philosophical Transactions of the Royal Society B*, *377*(1866), 20210333.

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

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.

Sim, Z. L., & Xu, F. (2017). Infants preferentially approach and explore the unexpected. *British Journal of Developmental Psychology*, *35*(4), 596-608.

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, *348*(6230), 91-94.

Stahl, A. E., & Feigenson, L. (2024). Young children distinguish the impossible from the merely improbable. *Proceedings of the National Academy of Sciences*, *121*(46), e2411297121.

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*, *104*(48), 19156-19159.

Téglás, E., Ibanez-Lillo, A., Costa, A., & Bonatti, L. L. (2015). Numerical representations and intuitions of probabilities at 12 months. *Developmental Science, 18*(2), 183-193.

Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science*, *332*(6033), 1054-1059.

Xu, F., & Denison, S. (2009). Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition*, *112*(1), 97-104.

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences*, *105*(13), 5012- 5015.