PNAS BRIEF REPORT | PSYCHOLOGICAL AND COGNITIVE SCIENCES OPEN ACCESS

Check for updates

Young children distinguish the impossible from the merely improbable

Aimee E. Stahl[a,1] and Lisa Feigenson[b]

Edited by Renée Baillargeon, University of Illinois at Urbana-Champaign, Champaign, IL; received June 10, 2024; accepted September 9, 2024

From infancy, children show heightened interest in events that are impossible or improbable, relative to likely events. Do young children represent impossible and improbable events as points on a continuum of possibility, or do they instead treat them as categorically distinct? Here, we compared 2- and 3-y-old children's learning (N = 335) following nearly identical events that were equi-probable, improbable, or impossible. We found that children learned significantly better following impossible than possible events, no matter how unlikely. We conclude that young children distinguish between the impossible and the merely improbable.

surprise | learning | possibility | children

People evaluate the likelihood of events from their earliest days. Even infants notice events that are surprising either because they are physically impossible [e.g., a ball passing through a wall; (1)] or statistically improbable [e.g., a random sample of white balls drawn from a population of mostly red balls; (2, 3)].

But are impossible events represented differently from improbable ones? Many models of human reasoning treat impossible events and highly improbable events as nearby locations on a continuous scale, making no categorical distinction between them (4, 5). Past work finds that school-age children treat improbable events as impossible (6, 7), at least when they are presented with verbal event descriptions and must make explicit possibility judgments, suggesting that they draw no bright line between the two. But to date, it is unclear whether young children distinguish the impossible from the improbable.

Here, we indexed children's sensitivity to possibilities using an intuitive task involving no explicit judgments. We capitalized on findings that infants and older children learn better following impossible than possible events (8, 9), and measured how well 2- and 3-y-old children learned following equi-probable, merely improbable, and impossible events that were otherwise identical. We asked whether children would show a graded learning enhancement, consistent with the lack of an impossible/improbable distinction, or a categorical learning enhancement, consistent with representing a clear distinction between the impossible and the improbable.

Methods

Two- and 3-y-old children (N = 335) saw a gumball-machine containing a visible population of novel objects; inserting a coin caused one object to "randomly" emerge through a chute. In the Equi-Probable condition (n = 64), the machine was seen to contain a mix of two types of novel objects: 10 novel objects of type A intermixed with 10 of type B–thus, there was a 50% chance of randomly receiving a type B object. In the merely Improbable conditions, the machine contained a majority of type A objects. In the 10% Probability condition (n = 26), there were 18 type A and 2 type B objects; in the 5% Probability condition (n = 64), there were 19 type A objects and 1 type B object; in the 2.5% Probability condition (n = 35), there were 39 type A objects and 1 type B object. These were chosen because even much younger infants are surprised when the rare object is drawn from similar populations (2). In the Equi-Probable and Improbable conditions, the experimenter pointed to the machine and said, "There's some purple and some pink, and that's all that's in there!"

For children in the Impossible condition (two independent samples: n = 63 and n = 83), the machine contained 20 objects of type A–there was a 0% chance of receiving a type B object. In the Impossible condition, the experimenter pointed to the machine and said, "There's some purple, and that's all that's in there!"

All children inserted a coin and received a Type B object (surreptitiously preloaded into the chute), which the experimenter immediately labeled as a *blick*. In the single test trial, children saw the type B object and two distractor objects that had also been given novel labels in prior filler trials (*SI Appendix*), and were asked to find the *blick*. The key measure was how well children learned this last novel word.

Results

First, we confirmed that there were no learning differences between the two samples of children in the Impossible condition (62% correct and 59% correct), $P = 0.74$; these were

Author affiliations: [a]Department of Psychology, The College of New Jersey, Ewing, NJ 08618; and [b]Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD 21218

Author contributions: A.E.S. and L.F. designed research; A.E.S. performed research; A.E.S. analyzed data; and A.E.S. and L.F. wrote the paper.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: stahla@tcnj.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2411297121/-/DCSupplemental.

Published November 4, 2024.

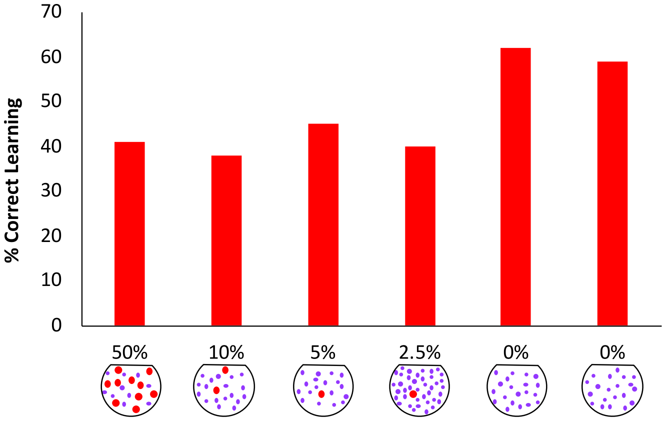

Fig. 1. Percent of children correctly identifying the Type B object when asked for the *blick* in the Equi-Probable, Improbable, and Impossible conditions. *X*-axis indicates probability of drawing a Type B object across conditions.

then combined into a single Impossible group. Children's learning across the Improbable subconditions also did not differ: 10% Probability (38% correct), 5% Probability (45% correct), and 2.5% Probability (40% correct), $\chi^2(2) = 0.47$, $P = 0.79$, so we combined these for subsequent analyses (Fig. 1).

Next, we compared children's learning in the Equi-Probable (41% correct), Improbable (42% correct), and Impossible (60% correct) conditions and found that these differed, $\chi^2(2) = 11.30$, $P = 0.004$. Post hoc two-tailed Fisher's exact tests, Bonferroni-corrected for multiple comparisons (adjusted alpha of 0.017), revealed that children learned significantly better following Impossible than Equi-Probable ($P = 0.01$) and Improbable outcomes ($P = 0.004$), and their learning was equivalent following Equi-Probable and Improbable outcomes ($P = 0.88$).

Finally, children performed similarly regardless of age; two-tailed Fisher's exact tests revealed no performance differences between 2- and 3-y-olds in the Equi-Probable ($P = 0.80$), Improbable ($P = 0.72$), or Impossible conditions ($P = 0.86$).

Discussion

Some approaches to characterizing human cognition suggest that all past and future events are represented with graded certainty (5). In line with this view, older children have failed to consistently distinguish highly improbable from impossible events (6, 7). In one experiment, 4-y-olds showed a mixed response pattern when making explicit possibility judgments—they successfully affirmed that, unlike impossible events, some probable *and* improbable events "could" happen, but they also incorrectly denied that other improbable events could happen (6). Here, we find that children do represent the impossible and the improbable as categorically distinct in situations involving a highly visualizable space of possible outcomes, such as when reasoning about the populations of colored objects used here and in ref. 6. Children in our task made this distinction spontaneously, as young as 2 y of age, without ever being asked about possibilities.

How did children reason about the events in our study? It could be that they engaged in modal reasoning, deploying symbolic representations that explicitly marked states of the world as possible, probable, impossible, or necessary. In the Equi-Probable and Improbable conditions, children might have thought, "The toy could be purple *or* it could be pink"; in the Impossible condition they might have thought, "It must be purple." These are modal constructs. Alternatively, children might have relied on nonmodal representations. One proposal is that "minimal representations of possibility" account for children's behavior—in uncertain situations, they simulate a single future outcome, randomly drawn from the space of possible outcomes, and represent this as actual as opposed to only possible (10). If the simulated and observed outcomes mismatch, children are surprised. The most straightforward version of this account struggles to capture our finding of nongraded learning, because as the likelihood of drawing a type B object gradually decreased from 50 to 2.5%, children would be decreasingly likely to randomly simulate the type B outcome, and therefore, surprise should gradually increase across these conditions (10); this is not the pattern we observed. Modifications to this account might be able to capture our results. Rather than randomly sampling a future state, perhaps children observe the machine's outcome and then backward-simulate the object's history. In the Impossible condition, no simulated sequence could have generated the outcome, potentially leading to surprise and enhanced learning. Or perhaps, using the experimenter's verbal descriptions of the machine's contents, children could represent an impossible event as "[not] × state of the world," thereby preventing it from being simulated. Future work is needed to adjudicate among these as candidate representations of impossible events.

Finally, our findings also suggest that although improbable events are surprising, even for infants (2), surprise itself does not automatically or directly enhance learning (8, 9). We propose that surprising events enhance learning specifically when they lack plausible explanations (11–13). Explanations, which allow revision of an initial model of the world to account for unexpected observations, are very likely to be sought following impossible events. Consider drawing an object from a population containing none of that object type (as in our Impossible condition). To maintain coherent commitments about objects and their causal interactions, model revision is required—maybe the object had dynamic features that altered its appearance, or the machine had a hidden compartment, or the experimenter played a trick. Indeed, we predict that enhanced learning will occur for any aspect of the event related to significant model revision (14). In contrast, merely improbable events do not always force model revision. Drawing one rare object from among 39 frequent objects is very improbable, but the observer need not revise their model of the objects or the machine to explain this outcome, because the scene statistics and machine mechanism already offer a satisfactory account. Still, we suggest that some improbable events will induce explanation-seeking. For example, seeing someone do something possible but unexpected (like taking a circuitous path to a goal) might impel the observer to revise their model of the person or event to accommodate candidate explanations for the behavior. In such cases, learning should be enhanced.

Returning to the question we began with: Are impossible events represented as distinct from improbable events? Our results show that although highly improbable events may at times seem similar to impossible events (15), even young children distinguish between them.

Materials and Methods

Participants were 23 to 49 mo old (M = 36 mo, 15 d; range = 23 mo, 1 d to 48 mo, 20 d; 158 girls). This study was approved by The College of New Jersey and the Johns Hopkins Homewood Internal Review Boards. Parents gave written informed consent. Methodological details appear in *SI Appendix*.

Data, Materials, and Software Availability. Anonymized data have been deposited in Harvard Dataverse (DOI: 10.7910/DVN/ZQ1PF9) (16). All other data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. We thank N. Chaudhari, C. Heiman, S. Mandaloju, C. Peters, and A. Smith-Flores for help with data collection.

1. E. S. Spelke, K. Breinlinger, J. Macomber, K. Jacobson, Origins of knowledge. *Psychol. Rev.* **99**, 605 (1992).
2. F. Xu, V. Garcia, Intuitive statistics by 8-month-old infants. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 5012–5015 (2008).
3. E. Téglás, V. Girotto, M. Gonzalez, L. L. Bonatti, Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19156–19159 (2007).
4. E. Borel, *Probabilities and Life* (Dover, 1962), Translated by M. Baudin.
5. M. Oaksford, N. Chater, *Bayesian Rationality: The Probabilistic Approach to Human Reasoning* (Oxford University Press, 2007).
6. A. Shtulman, S. Carey, Improbable or impossible? How children reason about the possibility of extraordinary events. *Child Dev.* **78**, 1015–1032 (2007).
7. A. Shtulman, B. Phillips, Differentiating "could" from "should": Developmental changes in modal cognition. *J. Exp. Child Psychol.* **165**, 161–182 (2018).
8. A. E. Stahl, L. Feigenson, Observing the unexpected enhances infants' learning and exploration. *Science* **348**, 91–94 (2015).
9. A. E. Stahl, L. Feigenson, Expectancy violations promote learning in young children. *Cognition* **163**, 1–14 (2017).
10. B. P. Leahy, S. E. Carey, The acquisition of modal concepts. *Trends Cogn. Sci.* **24**, 65–78 (2020).
11. M. I. Foster, M. T. Keane, M.T., Why some surprises are more surprising than others: Surprise as a metacognitive sense of explanatory difficulty. *Cogn. Psychol.* **81**, 74–116 (2015).
12. F. Margoni, L. Surian, R. Baillargeon, The violation-of-expectation paradigm: A conceptual overview. *Psychol. Rev.* **131**, 716–748 (2023).
13. J. Perez, L. Feigenson, Violations of expectation trigger infants to search for explanations. *Cognition* **218**, 104942 (2022).
14. A. E. Stahl, L. Woods, Infants preferentially learn from surprising teachers. *Infancy* **27**, 887–899 (2022).
15. K. Halvor Teigen, M. Juanchich, A. H. Riege, Improbable outcomes: Infrequent or extraordinary? *Cognition* **127**, 119–139 (2023).
16. A. E. Stahl, L. Feigenson, Data from "Young children distinguish the impossible from the merely improbable." Harvard Dataverse. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ZQ1PF9. Deposited 18 July 2024.

PNAS

Supporting Information for
Young Children Distinguish the Impossible from the Merely Improbable

Aimee E. Stahl* and Lisa Feigenson
The College of New Jersey and Johns Hopkins University
* Aimee E. Stahl.
Email: stahla@tcnj.edu

This PDF file includes:

Supporting text
SI References

Other supporting materials for this manuscript include the following:

Dataset S1

Supporting Information Text

Method

All children first were presented with two Filler Trials that were novel but possible, followed by the critical Probability Trial that was either Equi-Probable, Improbable, or Impossible. At the end of each of these (Filler Trial 1, Filler Trial 2, Probability Trial), a novel object involved in the event was labeled with a novel word (see 1). In the Test Trial we then examined whether children had learned the name of the novel object from the Probability Trial.

Participants. One hundred seventy-eight participants were tested in a university child development laboratory and 157 in the children's wing of a science museum. Performance did not differ by testing location. Thirty-seven additional children were excluded from analysis for being distracted or not following instructions (21), experimenter error (4), parent or sibling interference (5), or having first watched another participant complete the study in the museum (7). Children received a small prize for their participation.

Stimuli. A plush puppy (15 cm high) was used to engage children in the game. The first Filler Trial used bell-shaped cups with a blue checkerboard pattern (9cm high, 14.5cm diameter), and the novel object was a blue translucent plastic puck (2.5 cm high, 5cm diameter). The second Filler Trial used a green pencil box (3cm x 16cm x 11cm), and the novel object was an orange plastic lip-shaped whistle (4cm x 6cm). The Probability Trial used a clear gumball machine (16cm x 16cm x 10cm) that held novel objects. On top was a black foam-core box (9cm x 12cm x 9.5cm) with a coin slot. Attached to the machine was a faux chute whose front opening was covered with a decorated door; hidden in the back was a black felt flap that concealed an object behind it, such that the chute appeared empty. The novel objects in the Probability Trial were purple foam spirals (7cm high, 4cm diameter) and pink plastic barrel-shaped hair curlers (7.5cm high, 3cm diameter), except in the 2.5% Probability Trial where objects were cut in half to allow them to fit in the machine. In the Improbable conditions, the Type B object(s) were placed to ensure they were visible to children, and never in immediate proximity to the chute. For example, the singleton Type B object in the 5% and 2.5% Probability conditions was placed in the center of the machine toward the front, clearly visible but partially occluded by Type A objects. This ensured that the Type B object would not be expected to emerge from the chute based on spatial positioning (2).

Procedure. The experimenter introduced children to Puppy and said she needed help finding toys.

Filler Trials. In the first Filler Trial, the experimenter showed children the novel puck-shaped object under one of the two overturned cups. The experimenter placed a hand on each cup and said, "Watch this!" as she swapped their locations. She lifted the cup under which the object was hidden, pointed, and said, "Look what we got! This is a *dax!*" In the second Filler Trial, the experimenter placed the pencil box on the table and pretended to struggle to open it. The experimenter then asked the child to help. When they opened the box, the experimenter removed the object inside, pointed, and said, "Look what we got! This is a *fep*!"

Probability Trial. The experimenter placed the gumball machine on the table and said, "Do you see this machine I have? Do you see everything that's in there?" In the Equi-Probable and Improbable conditions (50%, 10%, 5%, 2.5%), where the machine contained purple spirals and at least one pink hair curler, she said, "There's some purple and some pink, and that's all that's in there!" In the Impossible condition, where the machine contained only 20 purple spirals, she said, "There's some purple, and that's all that's in there!" The experimenter then told the child that the machine worked just like a gumball machine, and that when a coin was inserted, one toy would come out. The experimenter then said, "We don't know which toy it will be!" She turned the machine around and opened the door to the chute to show that it was empty, saying, "See, nothing in there yet!" In all conditions, there was actually a pink toy preloaded behind the hidden flap in the chute. The experimenter then placed Puppy directly in front of the machine, saying, "Let's put Puppy right here so she can see everything!" This was done so that Puppy blocked most of the child's view of the machine's contents (since the machine was not actually functional and no objects shifted position when the coin was inserted). The experimenter then said, “Now remember, when we put our special coin in, just one toy is going to come out, and we don’t know which toy it will be.” The experimenter handed the child the coin to insert into the machine.

The experimenter then said, “Let’s see what we got” as she reached into the chute and retrieved the pre-loaded pink toy, which she pointed to and said, “Look what we got! This is a *blick*!” She then removed everything from the table.

Test Trial. The experimenter showed children a tray containing the three novel objects (blue puck from Filler Trial 1, orange whistle from Filler Trial 2, pink curler from the Probability Trial, with object position counterbalanced across children), and said, “Can you show me which one is the *blick*?” If children did not answer, the experimenter repeated the question up to two times. If a child did not answer the question after three attempts, which was rare, they were coded as answering incorrectly. Each child received a single test trial. The number of 2- and 3-year-olds who answered the test trial correctly in the Equi-Probable, Improbable, and Impossible conditions can be found in Dataset S1.

Dataset S1 (separate file). Dataset can be found at https://doi.org/10.7910/DVN/ZQ1PF9.

SI References

1. A. E. Stahl, L. Feigenson. Expectancy violations promote learning in young children. *Cognition*, **163**, 1-14 (2017).
2. E. Téglás, E. Vul, V. Girotto, M. Gonzalez, J. B. Tenenbaum, L. L. Bonatti. Pure reasoning in 12-month-old infants as probabilistic inference. *Science,* **332,** 1054-1059.