Goldilocks pattern of learning after observing unexpected physical events

***Qiong Cao[1,2] (qcao246@mit.edu), *Di Liu[1] (dliu88@jhu.edu), and Lisa Feigenson[1] (feigenson@jhu.edu)**
[1]Department of Psychological & Brain Sciences, Johns Hopkins University
[2]Department of Brain and Cognitive Sciences, MIT

Abstract

Infants learn better following expectancy violations. Yet it is unknown whether this surprise-induced learning operates across development, is all-or-none or graded, and whether surprise directly mediates it. We addressed these questions by showing adults events depicting varying numbers of violations. In Experiments 1 and 2, adults saw events with 0 to 3 physical violations, then heard a novel verb for the presented action. Adults learned better after observing violations; notably, their learning exhibited a Goldilocks pattern—initially increasing with number of observed violations, then declining. Experiment 3 asked whether this learning enhancement was driven by surprise itself, or by the search for explanations for the surprising events. Adults saw events with different numbers of violations, then rated their surprise and generated candidate explanations. Whereas surprise increased monotonically with violations, explanation-generation exhibited a Goldilocks pattern like that in Experiments 1-2. This suggests that surprise-induced learning may reflect the search for explanations.

Keywords: surprise-induced learning, surprise, learning, explanation, physical violation

Introduction

From early in life, infants have expectations about the physical and social world. These expectations have been revealed in several decades of violation of expectation studies, which find that infants look longer at events that violate adults' expectations about objects or agents, compared to events that accord with those expectations (Spelke, 2022; Spelke & Kinzler, 2007).

In addition, recent work finds that infants not only detect violations of object behavior, but also learn more from them compared to closely matched events without violations. For example, after seeing an event that violated physical expectations (e.g., an object appearing to pass through a wall), infants exhibited enhanced learning about the object, compared to seeing no violation (Stahl & Feigenson, 2015). This surprise-induced learning effect has also been shown in toddlers and early school-age children (Stahl & Feigenson, 2017, 2024), in the social domain (Cao et al., in revision), and across different testing platforms (Smith-Flores et al., 2021).

However, several aspects of surprise-induced learning remain uncharacterized. First, it is unknown whether surprise enhances learning beyond childhood, in mature learners. On the one hand, targeting events that defy expectations as special learning opportunities would be a rational approach for learners of any age; indeed, research on responses to prediction errors by human adults and non-human animals suggests that expectancy violations can improve their learning (e.g., Den Ouden et al., 2012; Holland & Schiffino, 2016). However, this prediction error literature has focused on violations of recently acquired, arbitrary contingencies, typically of specific stimulus-reward relationships, leaving open whether violations of the ”core knowledge” that guides everyday interactions with objects and agents also impact adult observers.

Second, is surprise-induced learning an all-or-none or a graded phenomenon? Past research using the violation of expectation paradigm has employed almost exclusively binary manipulations, comparing infants' responses to the presence versus absence of an expectancy violation (for an exception, see Téglás et al., 2011). The same is true for studies of early surprise-induced learning (Cao et al., in revision; Smith-Flores et al., 2021; Stahl & Feigenson, 2015, 2017). But might greater amounts of surprise promote greater amounts of learning? Past research suggests that adults view some surprising events as *more surprising* than others (Lewry et al., 2021; McCoy & Ullman, 2019); however, the question of whether the degree to which an observer's expectations are violated determines their learning remains unanswered. One possibility is that learning increases monotonically as a function of surprise. Alternatively, surprising might initially enhance learning, but when an event is too surprising, learning might actually decrease—that is, surprise-induced learning might exhibit a Goldilocks pattern (c.f., Kidd et al. 2012).

Finally, the mechanism underlying surprise-induced learning is not yet understood. It could be that learning is driven directly by the experience of surprise, such that feeling more surprised straightforwardly leads to stronger learning. Alternatively, learning may not be driven by surprise per se, but by related cognitive processes, such as explanation-seeking. On this account, surprise can trigger the observer to seek an explanation for what happened, so they may align their mental model to account for the unexpected observation(s). This search for an explanation in turn may lead the observer to gather additional information—a search reflected in enhanced learning about the entities involved in the surprising event. One piece of evidence supporting this explanation-based account comes from Perez and Feigenson (2022), who found that infants stopped exploring a surprising object once an explanation for the observed violation was offered.

To address the above questions about the nature of surprise-induced learning, we first asked whether adults exhibit

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

surprise-induced learning at all, and if so, whether it is all-or-none or graded. In Experiments 1 and 2, adults saw events in which an object's behavior either violated or accorded with principles of intuitive physics. Critically, we varied the number of violations adults saw (either cumulatively across trials in Experiment 1, or simultaneously in Experiment 2), so that on each trial they saw an object commit either zero, one, two, or three violations of typical behavior. Following each event, adults were taught a novel verb for the object's behavior, and then later were tested on their learning. Our main questions were whether adults would learn more from surprising than expected events, and whether the number of surprising events they saw would modulate their learning. Next, in Experiment 3, we asked whether adults' surprise-induced learning was driven by surprise or by explanation-seeking. Adults saw events similar to those in Experiment 2, and then were asked to rate how surprising they found the events, and to provide explanations for how the events could have occurred.

To preview, we found that (1) adults, like young learners, showed enhanced learning following expectancy violations, (2) this surprise-induced learning exhibited a Goldilocks pattern whereby it initially increased along with the number of observed violations, then diminished when too many violations were observed, and (3) the learning pattern seen in Experiments 1 and 2 was better aligned with observers' explanation-generation abilities than with their surprise ratings. This suggests that surprise-induced learning likely reflects the underlying drive to explain surprising events.

Experiment 1

Methods

Participants Undergraduate students ($N = 108$) participated in exchange for extra course credit.

Stimuli and Procedure Participants watched a sequence of videos presented online on their computers. First they saw three videos in which Object A engaged in three novel but physically possible actions (e.g., one action involved being repeatedly placed into and removed from a bucket). After each event, a female speaker labeled the action with a novel verb (e.g., "Look! It got [daxed]! Yeah, it got [daxed]!").

Next, participants saw three *key videos*, all involving Object B. Half of the participants were assigned to the *Violation of Expectation Condition*, in which Object B violated a physical principle in each key video: it appeared to pass through a solid wall (solidity violation), then was hidden in one location but retrieved from a different location (spatiotemporal continuity violation), then was pushed off a supporting surface but floated in mid-air (support violation). The other half of participants were assigned to the *No Violation Condition*; they saw Object B perform perceptually similar actions but always adhere to physical principles (e.g., it was stopped by the wall, was retrieved from the location where it had been hidden, and remained fully supported). The action in each event was then labeled with a novel verb. Event order (solidity, continuity, support) was randomized across participants.

Finally, all participants saw three more videos in which Object C participated in three novel but possible events. Thus in total, each participant saw 9 actions and learned 9 novel verbs (biffed, bindled, chayded, daxed, mipped, prammed, ravved, sigged, and torped), with verb-action pairings consistent across participants.

After watching all 9 videos, participants were tested on their learning of only the three novel verbs that had been presented in the key videos; these were tested in random order. On each test trial, participants saw three images and selected the one that corresponded to a prompted verb that had been taught earlier (e.g., "Earlier you saw an object get [daxed]. Which of these pictures shows that?") (Figure 1). Participants received no feedback.

Results

First we asked whether adults exhibited surprise-induced learning. We compared average learning scores in the Violation Condition and No Violation Condition; no overall difference was observed ($F(1,105) = 1.25$, $p = .267$). However, analysis of learning across the first, second, and third events within the test block revealed important differences. Participants' learning of the novel verb was no better after seeing an initial violation, compared to no violation, ($t(104.7) = 1.06$, $p = .291$), but learning was significantly better following the second violation, compared to the second non-violation counterpart ($t(104.4) = 2.04$, $p = .044$). Learning scores of the action shown in the third key event again did not differ across the Violation and No Violation conditions ($t(104.5) = -.65$, $p = .520$). Thus, adults did experience surprise-induced learning—an effect that emerged the second time they saw an object defy expectations.

Next we asked whether this learning enhancement was all-or-none or graded. First we calculated the mean learning performance across the first, second, and third trials in the No Violation condition, to establish a learning baseline in the absence of any violations (because in the No Violation condition, the number of violations that object B committed was zero across all trials). For trials in the Violation condition, the number of violations represented either the first, second, or third time the object had behaved surprisingly. When we compared learning after zero, one, two, or three violations had been seen (see Figure 2(I)), we observed that learning initially increased with the number of violations seen, and then decreased, exhibiting a Goldilocks pattern. To test this effect statistically, we first ran a linear mixed-effects model examining learning performance across the 4 violation types (none, first, second, third); this yielded a marginally significant omnibus effect of number of violation, $\chi^2(3) = 7.71$, $p = .052$. Post-hoc t-tests showed that participants learned significantly better following the second object violation compared to the second time no violation had occurred ($t(268) = 2.02$, $p = .044$) and compared to the third violation ($t(210) = 2.25$, $p = .025$). A Jonckheere-Terpstra test for monotonicity provided converging evidence that adults' learning performance did not change monotonically as the

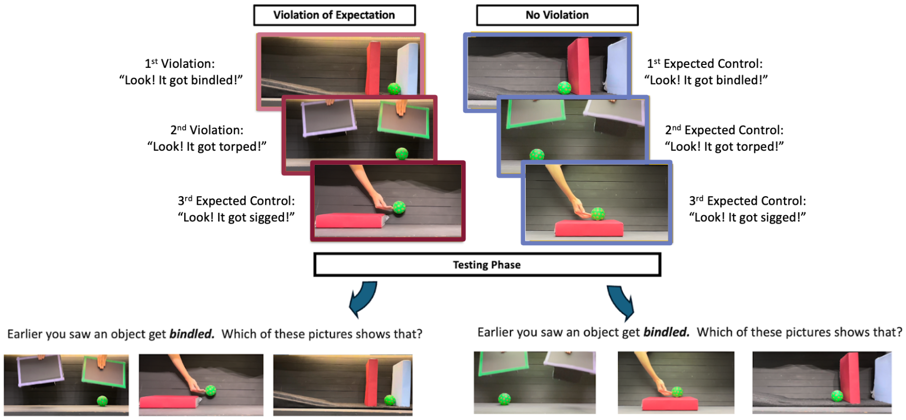

Figure 1: Conceptual Diagram of Experimental Procedure in Experiment 1

number of violations increased ($JT = 18034$, $p = .242$).

Discussion

In Experiment 1, we found that adults, like infants and young children, exhibit surprise-induced learning; participants learned better after successively seeing two violations compared to learning when no violations had been observed. In addition, we found evidence hinting that surprise-induced learning is not an all-or-none effect in adults. Instead, it showed a graded pattern as objects engaged in increasing numbers of surprising events. Notably, this graded learning followed a Goldilocks (c.f., Kidd et al., 2012)—rather than a monotonic—pattern, initially increasing with the number of observed violations before diminishing when the object violated expectations more than twice.

In Experiment 1, we manipulated expectancy violations by changing the history of the target object, modulating the number of times the object had already violated expectations prior to the moment of teaching. A different way of manipulating the degree of violations, however, is to adjust the extent to which a *single* event deviates from the observer's expectations. For example, an object can defy multiple expectations all at once—it can be seen to both float in midair and pass through a solid barrier simultaneously. Next, in Experiment 2, we asked whether experiencing different numbers of simultaneous violations would also affect learning. This served as a conceptual replication of Experiment 1.

Experiment 2

Methods

Participants Participants were 320 undergraduate students who participated in exchange for extra course credit.

Stimuli and Procedure Participants watched a sequence of 18 videos presented online, each depicting a different novel event that was then immediately labeled with a novel verb. Each sequence contained four key videos and 14 filler videos.

The four key videos depicted a simple physical event in which an object was placed at the top of a ramp and then released, so that it rolled down the ramp and passed behind an occluding screen above which a salient red wall protruded (Figure 3). The occluding screen was then removed, revealing one of four outcomes. On *No Violation* trials, the object was shown to have been stopped by the red wall in its path. On *Single Violation* trials, the object violated one of four types of expectations, committing either: a solidity violation (the object was now on the far side of the wall, as though it had passed through it), a feature violation (the object changed its color), a kind violation (the object appeared to have changed form, e.g., a ball changed into a similarly colored car), or a support violation (the object appeared to float in midair). On *Double Violation* trials, the object committed two of the above violations (e.g, it appeared to have passed through the wall and to have changed from a ball to a car). On *Triple Violation* trials, the object committed three of the above violations (e.g., it appeared to have passed through the wall, changed color, and float in mid-air). Each key video started with a unique object; which starting object appeared in which type of key trial was randomized across participants. The 14 filler videos each depicted a simple physical event that was novel but did not violate any physical principles (e.g., an object was transferred back and forth between two containers; an object was tapped with plastic rods). As in Experiment 1, immediately after the event ended (whether in key trials or filler trials), the depicted action was labeled with a novel verb ("Look! It [daxed]! Yeah, it [daxed]!") After the first 9 trials and again after the second 9 trials, participants received three test trials in which three images appeared and they were prompted to select the one corresponding to one of the verbs taught earlier ("Which one [daxed]?"). No feedback was given. We initially tested 102 participants using this design, with the order of key videos, verb-event pairings, and image positions on each test trial pseudo-randomized. Later

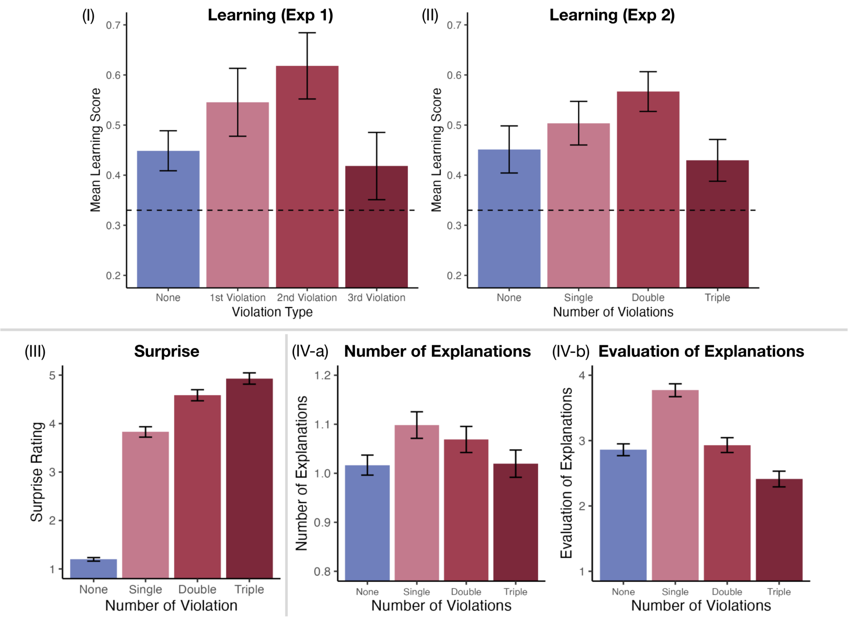

Figure 2: (I-II) Learning scores plotted against the number of violations participants observed sequentially (I, Exp 1) or simultaneously (II, Exp 2); (III) Surprise ratings plotted against the number of violations participants observed; (IV-a) Number of explanations generated, and (IV-b) the quality of explanations generated.

we tested 218 additional participants using a fully randomized design, and including two attention-check questions, designed to filter out inattentive participants.

For participants in the experiment version that included attention-check questions, we only analyzed responses from test blocks in which they had correctly answered the attention-checks. Preliminary analyses revealed that the data from the two versions of the experiment did not differ, so we combined them for further analysis. Finally, to minimize the influence of any top-down learning strategies, we restricted our analyses to the first half of the experiment, which included two key trials from each participant.

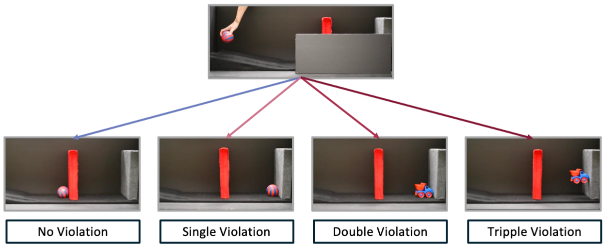

Figure 3: Examples of Events Presented in Experiment 2

Results

To test for the presence of surprise-induced learning, we compared participants' learning scores following events with no violations to those with one or more violations. This revealed no significant difference overall ($\beta = 0.06$, $SE = 0.05$, $p = .284$). However, as in Experiment 1, we observed significantly enhanced learning following events containing a double violation, compared to events containing no violations ($t(431) = 2.00$, $p = .046$), but not in other comparisons ($ps > .387$). Again similarly to Experiment 1, as the number of observed violations increased, participants' learning performance first increased and then decreased (Figure 2(II)). A Jonckheere-Terpstra test provided evidence that adults' learning performance did not increase monotonically with number of violations ($JT = 55111$, $p = .563$).

Discussion

In Experiment 2, we tested adults' learning of novel words following different numbers of violations of object behavior. We replicated the pattern found in Experiment 1: adults exhibited a Goldilocks pattern of learning, whereby they tended to learn better following events that violated two aspects of object behavior, but when more violations occurred, their learning declined. Thus, the findings from Experiment 2 conceptually replicate those of Experiment 1, suggesting that the Goldilocks pattern of surprise-induced learning obtains regardless of whether expectancy violations accumulate over time (Experiment 1) or occur simultaneously (Experiment 2).

What processes underlie this learning enhancement? As described above, there are at least two candidates: surprise could directly drive learning, or learning could be amplified by the drive to find an explanation for the observed event. In Experiment 3 we sought evidence to help decide between these possibilities. Participants saw videos similar to those in Experiment 2. This time, rather than measuring learning, we instead asked them to rate how surprising each video was, and to provide candidate explanations for the observed events. Our aim was to compare the overall shape of participants' surprise ratings and candidate explanations (as a function of number of violations) to the pattern of learning observed in Experiments 1 and 2.

Experiment 3

Methods

Participants Participants were 243 undergraduate students who participated for extra course credit.

Stimuli and Procedure Participants watched a sequence of 23 videos presented online. Eleven of these depicted an event that violated expectations about object behavior. The videos were identical to those in Experiment 2, and either depicted a Single Violation (of object solidity, featural continuity, kind continuity, or support), a Double Violation, or a Triple Violation. The remaining 12 videos were filler videos included to make the task sufficiently challenging; these never involved any violations of expectation. To avoid any influence of novel perceptual features from filler videos, and to encourage focusing on the key events, we replaced the filler videos from Experiment 2 with repetitions of the expected videos in which no violations took place. In these filler videos, an object simply rolled down a ramp behind an occluding screen; when the screen was removed, the object was revealed resting on the near side of a wall in its path. Each video lasted approximately 10 seconds.

Surprise measure. After each video, participants rated the degree to which they found the depicted event surprising ("How surprising did you find this event?"), using a 7-point scale from "1- not surprising" to "7- extremely surprising."

Explanation measure. To assess participants' ability to generate potential explanations for the stimulus events, for each participant we pseudo-randomly selected one No Violation video, one Single Violation video, one Double Violation video, and one Triple Violation video (we did not ask participants to generate explanations for all 11 videos, as pilot testing revealed effects of fatigue at being asked to type in large numbers of explanations). For these trials, after participants had entered their surprise rating, they were asked to explain how the event might have occurred ("Consider what happened in the video. Please type an explanation for how it happened").

The order in which Violation videos and filler videos were presented was pseudo-randomized so that videos with the same number of violations never appeared more than twice in a row. As in Experiment 2, we also included three attention-check trials to ensure that participants remained engaged and attentive; these involved responding to simple questions, such as reporting the number of people in an image.

Response Coding An experienced rater coded participants' explanations, quantifying the number of explanations provided on each trial and rating the degree to which participants' responses offered a plausible explanation for the observed event. For explanation number, if participants provided a single candidate account of the presented event, this was counted as one explanation. For instance, "there is a magnet on the wall" was counted as one explanation for a Single Violation event in which the object was floating. If participants offered multiple possible explanations for a violation, these were counted as such. For instance, "There could be a string attached to the car, or maybe there is a magnet" was counted as two explanations. Unlike explanations that referenced possible physical features of the array (such as invisible strings and holes in the wall), responses that referred to "magic" or video editing were coded as non-explanations, because our focus was on plausible explanations for actual live events (like those shown to infants). For explanation quality, participants' typed responses were rated from very poor (e.g., "I don't know" received a score of 0) to excellent (e.g., a specific account of the scene that was judged to fully explain the observed event received a score of 5). A second rater double-coded 50% of the responses; the intraclass coefficient between raters was .833 for the number of explanations provided, and .898 for the evaluation of the explanation's plausibility.

Results

Surprise ratings. We first examined participants' surprise ratings using a linear mixed effects model. As shown in Figure 2(III), we found that average surprise ratings for events depicting any number of violations of object behavior were significantly higher than ratings for events with no violations ($\beta = 3.06$, $SE = 0.03$, $p < .001$) As the number of violations increased from zero to three, participants' surprise ratings increased monotonically [$JT = 8673724$, $p < .001$). The difference between surprise ratings for No Violation, Single Violation, Double Violation, and Triple Violation was significant for all pairwise comparisons (ps $< .001$).

Explanations. We then asked whether the number of explanations participants generated was related to the number of violations they observed. We found that the number of explanations generated exhibited a Goldilocks pattern, with the greatest number of explanations generated for events containing a single violation of object behavior, and the fewest explanations generated for events containing three violations. We observed a marginally significant effect whereby participants generated more explanations after having seen any number of violations, compared to no violations ($\beta = 0.05$, $SE = 0.03$, $p = .058$). Participants generated significantly more explanations following Single Violations compared to No Violations ($t(343.47) = 2.63$, $p = .009$), but we found no significant difference in the number of explanations generated between Single and Double Violation trials, or between Double and Triple Violation trials (ps $> .252$).

Next we examined the quality of the explanations participants provided. As shown in Figure 2(IV), explanation quality exhibited a Goldilocks pattern as the number of violations increased, with participants' explanations receiving the highest quality rating for events containing a Single violation. We observed a marginally significant effect of better explanations when participants saw any number of violations, compared to seeing no violations ($\beta = 0.22$, $SE = 0.13$, $p = .080$). Pairwise comparisons yielded significant differences across each adjacent pair: No Violation vs. Single Violation ($t(378) = 6.67$, $p < .001$), Single Violation vs. Double Violation ($t(338) = -6.21$, $p < .001$), and Double Violation vs. Triple Violation ($t(321) = -3.60$, $p < .001$). We confirmed the non-monotonic trend with Jonckheere-Terpstra Test ($JT = 34272$, $p = 1.000$).

Discussion

In Experiment 3, we showed adults events containing different numbers of physical violations. We found that as the number of violations increased, adults reported feeling increasingly surprised. However, their explanation-related behaviors, namely the number of explanations they generated and the quality of their candidate explanations, showed a different pattern. Rather than increasing monotonically, they exhibited a Goldilocks pattern—producing the greatest number and highest quality of candidate explanations for events containing a single violation, with both measures declining thereafter.

Our goal in Experiment 3 was to ask what might be driving surprise-induced learning. Here, by comparing the way surprise-induced changed with the number of observed violations to the way surprise and explanation-generation changed with the number of violations, we took a first step towards identifying a potential mechanism. We found that both learning and explanation-generation exhibited a Goldilocks pattern, rather than the monotonic increase observed in participants' surprise ratings. This suggests that so-called "surprise-induced learning" may in fact be driven by observers' search for explanations, rather than directly by surprise itself.

General Discussion

In the current work we found that adults, like young learners, show enhanced learning after witnessing events that violate their expectations compared to those that align with them. In addition, we found that surprise-induced learning is not all-or-none but graded, in that it is sensitive to the number of violations experienced by the observer. Intriguingly, adults across two experiments exhibited a Goldilocks pattern whereby learning was initially enhanced by seeing a violation, but then appeared to diminish as the number of observed violations grew too large. Finally, this learning pattern did not align with that of self-reported surprise, which increased monotonically with the number of violations. Instead, it aligned with explanation-seeking behaviors: observers initially generated more explanations, and better explanations, after seeing an expectancy violation, but with too many violations, both explanation number and quality declined. These results suggest that, at least in adults, surprise-induced learning is better explained by an explanation-based account than by surprise alone.

These findings are consistent with prior research showing that adults are sensitive to expectancy violations. For example, previous work finds that adults rate events involving physical violations as significantly more surprising than events without violations (Smith et al., 2020). Our study extends this picture by revealing that adults also learn from such prediction errors. However, the nature of what can be learned following a violation remains ripe for future investigation. Here, we found that adults were better at learning the labels of surprising actions, whereas other work suggests that adults do not learn new physical rules from similar impossible events (Liu & Xu, 2022). One possibility is that enhanced learning occurs only for information that helps explain the surprising event. Although learning a new verb for a surprising action (as in Experiments 1 and 2) is not directly explanatory, words for objects and actions are often taken as referring to general classes rather than specific instances (Dewar & Xu, 2009). As such, learning a new word for an object (Smith-Flores et al., 2021) or action (Experiments 1-2 here) may be helpful as the observer tries to construct a mental model that can allow them to identify other instances where their prior predictions do not hold.

Our work also aligns with recent developmental research on explanation-seeking in infants. In one study, 11-month-olds who saw an object behave unexpectedly (e.g., a ball appeared to have passed through a wall), engaged in hypothesis-testing behaviors, such as banging the ball on the table to test its solidity, as if searching for an explanation for the unexpected event (Stahl & Feigenson, 2015). When infants were provided with an explanation—such as a hidden hole in the wall—their tendency to explore that initially surprising object disappeared (Perez & Feigenson, 2022). In the present work, we find that explanation-related behaviors following surprising events are graded, rather than all-or-none. Whether this is also true of infants, with their more limited world knowledge, remains an open question.

Importantly, the experiments reported here do not yet establish a causal link between explanation-seeking and learning (in part because different participants contributed to the data across our three experiments). In future work, we will ask whether providing adults with an explanation after they have seen an impossible event affects their learning. If learning is driven by explanation-seeking, then receiving an explanation should eliminate any surprise-induced learning.

In sum, here we offer evidence that adults exhibit a Goldilocks pattern of learning after observing different numbers of physical violations. We suggest that this graded learning is likely driven by explanation-seeking, which promotes successful mental model revision.

References

Cao, Q., Smith-Flores, A., Zhou, J. X., Perez, J., & Feigenson, L. (in revision). *Violations of social expectations enhance infants' learning*.

Den Ouden, H. E., Kok, P., & De Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in psychology*, *3*, 548.

Dewar, K., & Xu, F. (2009). Do early nouns refer to kinds or distinct shapes? evidence from 10-month-old infants. *Psychological Science*, *20*(2), 252–257.

Holland, P. C., & Schiffino, F. L. (2016). Mini-review: Prediction errors, attention and associative learning. *Neurobiology of learning and memory*, *131*, 207–215.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS one*, *7*(5), e36399.

Lewry, C., Curtis, K., Vasilyeva, N., Xu, F., & Griffiths, T. L. (2021). Intuitions about magic track the development of intuitive physics. *Cognition*, *214*, 104762.

Liu, R., & Xu, F. (2022). Can adults revise their core beliefs about objects? *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*(44).

McCoy, J., & Ullman, T. (2019). Judgments of effort for magical violations of intuitive physics. *PLoS One*, *14*(5), e0217513.

Perez, J., & Feigenson, L. (2022). Violations of expectation trigger infants to search for explanations. *Cognition*, *218*, 104942.

Smith, K. A., Mei, L., Yao, S., Wu, J., Spelke, E., Tenenbaum, J. B., & Ullman, T. D. (2020). The fine structure of surprise in intuitive physics: When, why, and how much? *Proceedings of the Annual Meeting of the Cognitive Science Society*, *42*.

Smith-Flores, A. S., Perez, J., Zhang, M. H., & Feigenson, L. (2021). Online measures of looking and learning in infancy. *Infancy*, *27*(1), 4–24.

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

Spelke, E., & Kinzler, K. D. (2007). Core knowledge. *Developmental science*, *10*(1), 89–96.

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

Stahl, A. E., & Feigenson, L. (2017). Expectancy violations promote learning in young children. *Cognition*, *163*, 1–14.

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., 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.