Violations of expectation trigger infants to search for explanations

Jasmin Perez*, Lisa Feigenson

Department of Psychological and Brain Sciences, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA

A R T I C L E   I N F O

Keywords:
Surprise
Explanation
Violation-of-expectation
Infants
Exploration

A B S T R A C T

Infants look longer and explore more following violations-of-expectation, but the reasons for these surprise-induced behaviors are unclear. One possibility is that expectancy violations heighten arousal generally, thereby increasing infants’ post-surprise activity. Another possibility is that infants’ exploration reflects the search for an explanation for the surprising event. We tested these alternatives in three experiments. First in Experiment 1 we confirmed that seeing an object violate expectations (by passing through a solid wall) increased infants’ exploration of the surprising object, relative to when no expectation violation was seen. Then in Experiment 2 we measured infants’ exploration after they had seen the same violation event, but then an explanation for the event was provided (the wall was revealed to have a large hole in it). We found that providing this explanation abolished infants’ surprise-induced exploration. In Experiment 3 we replicated this effect. Furthermore, we found that the longer infants looked at the explanation, the greater their reversal in exploratory preference (i.e., the more they ignored the surprising object). These findings demonstrate that preverbal infants both seek and recognize explanations for surprising events.

1. Introduction

A central finding in developmental psychology is that, by just a few months of age, infants look longer when objects behave in ways that adults describe as surprising or impossible, compared to ways that adults find unsurprising or predicted. Researchers have leveraged this methodological discovery to characterize pre-verbal expectations across a range of domains (see Spelke & Kinzler, 2007 for a review). As one example from the object domain, 2.5-month-old infants saw a ball roll down a ramp and pass behind a screen, heading towards a wall in its path. The screen was then lifted to reveal either the expected outcome of the ball at rest on the near side of the wall, as though the wall had interrupted its trajectory—or the unexpected outcome—the ball on the far side of the wall, as though it had passed straight through. Infants looked significantly longer at the unexpected outcome, suggesting that they expected the ball to be stopped by the wall (Spelke, Breinlinger, Macomber, & Jacobson, 1992). This implicates an early understanding of an important principle of object behavior: that objects respect solidity and cannot pass through one another. Other studies using the violation-of-expectation method reveal that infants have further expectations about objects, including that they continue to exist when hidden (Baillargeon & DeVos, 1991; Baillargeon, Spelke, & Wasserman, 1985), trace spatiotemporally continuous paths (Aguiar & Baillargeon, 1999; Wilcox, Nadel, & Rosser, 1996), and fall if unsupported (Needham & Baillargeon, 1993).

Violation-of-expectation findings also implicate early knowledge across several other domains, in addition to that of objects. To highlight a few examples, infants look longer when 5 objects + 5 objects yields 5 objects, compared to when 5 + 5 yields 10, pointing to core knowledge of numerosities (McCrink & Wynn, 2004). Infants look longer when someone reaches for a new object instead of an object for which they had previously reached, pointing to knowledge of people’s actions reflect their goals (Woodward, 1998). And infants look longer when someone approaches an agent who has previously been mean over an agent who has previously been nice, pointing to early knowledge of agents’ dispositions (Kuhlmeier, Wynn, & Bloom, 2003).

Results like these tell us that infants notice surprising events across a wide range of entities. But beyond simply detecting these events, more recent work also suggests that infants also experience changes in exploration and learning as a result of expectancy violations. In one study, 10- to 12-month-old infants saw a solidity event like that described above, and then were offered the opportunity to explore the object from the just-seen event as well as a novel distractor. Infants who had seen the object accord with their expectations (i.e., be stopped by the wall) generally ignored that object and spent most of their time exploring the distractor, consistent with a preference for novelty.
contrast, infants who had seen the object violate expectations (i.e., pass through the wall) mostly explored the familiar object that had just defied solidity. This shift in exploratory preference was also observed when infants saw violations of object support (an object appearing to float in mid-air) (Stahl & Feigenson, 2015). Similarly, Sim and Xu (2017) showed 13-month-old infants two boxes containing balls of many different colors, from which an experimenter “randomly” drew samples. The sample from one box yielded all identically-colored balls (a surprising outcome, given the population), and the sample from the other yielded multiple colors (an expected outcome, given the population). Later, infants spent more time touching and reaching into the box that had yielded the unexpected sample.

Thus, expectancy violations propel infants to look longer and explore more. However, the question of why violations-of-expectation have this effect—namely, whether these surprise-induced behaviors have any functional significance—remains unanswered. One possibility is that infants’ surprise-induced exploration reflects a targeted effort to learn from anomalous events. In particular, infants might explore because they are searching for explanations: infants could be examining objects in the service of gaining a better understanding of what just happened, actively seeking information that offers a causal mechanism for the violation just witnessed. Such targeted information-search would be an efficient way to learn about the world, giving infants the opportunity to revise their mental models when faced with contradictory evidence. Support for this view comes from the finding that 11-month-old infants’ post-surprise behaviors are linked to the type of object violation seen: infants who saw an object appear to float without support later tended to repeatedly drop the object, whereas infants who saw an object appear to pass through a solid wall repeatedly banged the object (Stahl & Feigenson, 2015). These diverging exploratory actions could reflect infants’ drive to explain the particular violation they had observed. However, it is also possible that this violation-action link simply reflects infants’ attempt to reproduce the surprising events, without seeking their causes.

There are other reasons to wonder whether infants’ surprise-induced exploration has anything to do with explanation-seeking. Surprising events might simply increase infants’ overall arousal, which could amplify their subsequent behaviors—including their interaction with objects. This possibility is consistent with findings that amphetamine-induced arousal increases motor activity in human adults (and mice) without increasing exploration per se (Minassian et al., 2016). One version of this general arousal account seems to be ruled out by the finding that surprise promotes infants to specifically explore the object involved in the surprising event, as opposed to unrelated distractor objects (Stahl & Feigenson, 2015). Still, it might be that arousal motivates infants to engage with familiar entities (e.g., the object from the preceding event), without any top-down goal of seeking specific information.

If infants are capable of seeking explanations for surprising events, then one would expect to find evidence that cognition cares about explanations from early in development. This appears to be the case by the toddler- to preschool years. By 2- to 5-years of age, children are already asking why and how questions in their everyday conversations with adults. Their questions are more persistent when they receive non-explanatory responses as opposed to explanations (Callanan & Oakes, 1992; Frazier, Gelman, & Wellman, 2009), and they prefer learning from people who provide non-circular over circular explanations (Corriveau & Kurkijärvi, 2014). Moreover, 3- to 5-year-old children can offer their own very explanatory, and accurate, explanations of events that they have seen in unexpected ways (Legare, Gelman, & Wellman, 2010). In addition to seeking and producing verbal explanations, preschool-aged children produce exploratory actions in situations where explanations are warranted. They explore more when it is clear that there is a new causal relation to be learned (Butler & Markman, 2012) and when faced with unexplained or ambiguous events; further, their exploratory behaviors seem targeted to potential explanations of the events (Schulz & Bonwitz, 2007). This ability to generate and test explanations in order to form a causal understanding of the world has been argued to be a central driver of children’s learning (Gopnik, 1998; Legare & Clegg, 2014; Liquin & Lombrozo, 2020; Wellman, 2011).

Yet, evidence that preverbal infants recognize explanations is surprisingly sparse. One potential indication that they might come from patterns of null results in some violation-of-expectation studies (Baillargeon, 1994). In one series of experiments, 3- and 3.5-month-old infants saw an object pass behind a U-shaped screen (Aguiar & Baillargeon, 2002). Sometimes the U was big enough so that its bottom portion hid the object as it moved from one side to the other, and other times the U was too short, such that the object should have been partially visible—but wasn’t. Three-month-old infants looked longer at the shorter-U event than the larger-U event, suggesting that they were surprised when the object passed by without being seen. Yet older, 3.5-month-olds failed to show this elevated looking. Baillargeon’s interpretation is that 3.5-month-olds spontaneously generated an explanation for the surprising event: they posited that two identical objects were present—one behind each arm of the U. The event would then be constructed as involving an object moving out from behind the left arm of the U and then, after a pause, a second object emerging from behind the right arm. Baillargeon’s account is bolstered by her additional finding that when 3.5-month-olds definitively saw that only one object was present prior to the event (because the U-shaped screen was lowered to reveal just a single object, then was raised again), they now looked longer in the shorter-U condition (just like the younger 3-month-olds who were not yet capable of generating an explanation)—potentially because the 2-object explanation could no longer account for their observations. Baillargeon (1994) suggests that other null results in violation-of-expectation studies could similarly be caused by infants constructing explanations—if infants faced with surprising events generate explanations on their own, then they may show no difference in looking time to impossible versus possible events (Baillargeon, 1994).

Although the findings described above could potentially implicates preverbal recognition of explanations, other evidence suggests that infants struggle to recognize explanations for surprising events. For example, Wang and Goldman (2016) found that 12-month-olds successfully used information that was presented before a surprising event to revise their expectations of the event’s outcome. Yet when the same information was presented after the event—thereby providing an explanation for what had already happened—infants seemed unable to integrate this new information into their event representation. The distinction between information presented before versus after a surprising event may be critical to thinking about whether infants genuinely represent explanations. When infants have all of the relevant information available to them before an event occurs, they may simply be triggered to represent the event in a different way. For example, in the task described above by Aguiar and Baillargeon (2002), infants in one condition saw the object move back and forth behind the small U-shaped screen. Although the object was tall enough to protrude above the horizontal portion of the U, it never did; it only emerged from and then disappeared behind the screen’s vertical arms. Baillargeon suggests that the 3.5-month-olds in this task formed an explanation for this seemingly impossible event: two objects were present, one behind each arm. This 2-object construal of the event can describe what happened because the 2-object explanation could no longer account for their observations. Baillargeon (1994) suggests that other null results in violation-of-expectation studies could similarly be caused by infants constructing explanations—if infants faced with surprising events generate explanations on their own, then they may show no difference in looking time to impossible versus possible events (Baillargeon, 1994).
then use subsequent information to resolve it. For this definition, the distinction between the ability to use information available before a surprising event (such that the event is never perceived as surprising) versus after a surprising event (to explain away the surprising conflict) is critical. With respect to this definition, it remains an open question whether infants recognize explanations.

Therefore, in the present work we sought to address the open question of whether infants can use new evidence to explain a previously observed violation-of-expectation event. We hypothesized that if surprise-induced exploration reflects the search for an explanation, then if offered a plausible explanation for the surprising event, infants’ post-surprise exploration should cease. We tested this hypothesis in three experiments. First, in Experiment 1, we attempted to replicate the finding that surprise propels infants to explore. We showed 11-month-old infants a physical event that either culminated in an expected outcome (that accorded with the principle of object solidity), or a nearly identical event that culminated in a surprising outcome (that violated the principle of solidity). Based on previous findings (Guillargeon et al., 1985; Spelke et al., 1992; Stahl & Feigenson, 2015), we anticipated that when then given the chance to explore the object from the event versus a novel distractor, infants who had seen the violation would show a heightened preference to explore the target object. Next, in Experiment 2, we asked what would happen if infants were offered an explanation for the surprising event. Infants saw the same surprising solidity violation as in Experiment 1, but then were given new information that offered a plausible explanation for what had just happened. If surprise simply increases infants’ arousal and motivation to engage with the object from the surprising event, then infants in Experiment 2 should show the same heightened exploration as infants in Experiment 1. Alternatively, if infants’ exploration reflects the search for an explanation for the anomalous event, then receiving an explanation should abolish any surprise-induced exploration. Finally, in Experiment 3, we replicated Experiment 2 with an independent sample of infants.

2. Experiment 1

Although hundreds of studies find that infants look longer at surprising than expected outcomes, to date only a few published studies report that surprise also causes infants to explore more (Sim & Xu, 2017; Stahl & Feigenson, 2015). Therefore, in Experiment 1 we first sought to replicate the finding that surprise enhances infants’ exploration. Eleven-month-old infants saw a solidity event modeled on that by Spelke et al. (1992), in which an object rolled down a ramp and passed behind a screen, heading towards a wall in its path. The event either culminated in an expected outcome (the object appeared to have been stopped by the wall) or a surprising outcome (the object appeared to have passed through the wall). We predicted that infants would explore the object more when it had been involved in the surprising event as opposed to the expected one.

2.1. Method

2.1.1. Participants

Forty-six full-term 11-month-old infants participated (mean age: 11 months, 22 days; range: 10 months, 12 days to 14 months, 1 day); twenty were girls. Data from nine additional infants were excluded for fussiness (3), parent interference (1), failure to look at the event outcome (4); see below), or failure to engage in any object exploration at all (1). Written informed consent was obtained from parents prior to participation, and infants received a small gift (book, stuffed toy) to thank them for their participation.

2.1.2. Stimuli

Infants saw two objects during the course of the study, chosen to be easy discriminable in basic level kind and low level visual features: a blue and red toy truck and a green and red ball. The truck measured 12 × 8 centimeters and the ball measured 8 centimeters in diameter. Half of the infants saw the truck serve as the Target Object and the ball as the Distractor Object, and half saw the reverse.

2.1.3. Procedure

Infants sat in a highchair one meter from a puppet stage. Parents sat out of view behind them and were instructed to remain silent throughout the study.

Familiarization Event. The study started with a trial designed to familiarize infants to the stage and the Target Object. The experimenter placed the Target Object (e.g., truck), and waved it for 5 seconds to draw infants’ attention before placing it at the top of the ramp. She said, “Watch this!” and released the object so that it rolled down the ramp and behind the occluding screen. The experimenter then said, “Look at this!” as she reached down and removed the screen, revealing the Target Object resting against the grey wall on the far right side of the stage. After five seconds, the experimenter lowered the stage curtain and the Familiarization trial was considered complete.

Solidity Event. All infants next saw a single test event. The experimenter raised the curtain to reveal the ramp, the far grey wall, and the black occluding screen, all in the same positions they had been in during the Familiarization event. In addition, there was now a bright red wall (35 cm high × 25 cm wide) interrupting the path from the ramp to the far grey wall; the top of this red wall could be clearly seen protruding 13 cm above the top of the occluding screen (Fig. 1). The experimenter drew infants’ attention to the visible portion of the red wall, saying “Look at this!” as she pointed to and tapped it for two seconds before withdrawing her hand from the stage. Next the experimenter reached back into the stage holding the Target Object, said, “Watch this!” and released the object at the top of the ramp so that it rolled down and disappeared behind the screen. After approximately five seconds she said, “Look at this!” as she removed the screen, revealing the final position of the Target Object. For infants in the Expected condition (N = 22), the Target Object was revealed on the near side of the red wall—consistent with the wall having stopped the object (Fig. 1, Panel 3A). For infants in the Surprising condition (N = 24), the Target Object was revealed on the far side of the red wall—consistent with the object having passed straight through (Fig. 1, Panel 3B). Infants’ looking at the outcome was measured for 10 seconds; looking was defined as visual fixation anywhere within the stage area, starting from the moment the black occluding screen was lifted out of view. Because we were interested in how infants’ exploration would change as a consequence of which outcome infants had seen, we excluded from analysis any infants who did not look at the outcome for at least 2 seconds (4 infants). After 10 seconds the experimenter said “Look at this!” as she picked up and rotated the red wall so infants could see its front face (which had previously not been visible). She placed the wall on the left side of the stage, facing infants, and infants’ looking at the wall and at the Target Object were measured for 20 seconds starting from the moment the wall was placed in its rotated position (Fig. 1, Panels 4A & 4B); looking at anything other than the wall was not coded. After 20 seconds, the experimenter lowered the stage curtain.

Exploration Phase. Immediately following the Test Event, the experimenter emerged from behind the stage holding the Target Object (e.g., truck) and a novel Distractor Object (e.g., ball), one in each hand. She said, “Look at these!” and placed both objects on the high-chair tray (with left-right placement counterbalanced across infants). Then she returned behind the stage so that she was out of infants’ view. The objects remained on the high chair tray for 30 seconds, during which
infants’ visual and manual exploration of the objects was recorded for later coding. As in previous work (Stahl & Feigenson, 2015), exploration was coded as the duration that infants looked at an object while touching it. All videos were coded offline by two trained observers. Each observer coded the videos of 60% of the participants; 20% of videos were coded by both observers. Inter-observer reliability was .93 for infants’ looking time and .95 for exploration.

2.2. Results

We first asked whether infants detected the solidity violation. We found that infants in the Surprising condition looked significantly longer at the event outcome ($M = 7.17$ sec, $SD = 2.51$ sec) than infants in the Expected condition ($M = 5.23$ sec, $SD = 1.71$ sec; $M_{\text{difference}} = 1.94, SEM = .64, t(44) = 3.04, p < .01, 95\% CI [.65, 3.23]$), confirming the results of many previous studies.

Next we asked whether the type of outcome infants saw affected their subsequent exploration. We first computed each infant’s exploration preference score by subtracting the time spent exploring the novel Distractor Object from the time spent exploring the familiar Target Object. Then we asked whether this exploration preference score differed across conditions. A 2 × 2 ANOVA with Condition (Surprising or Expected Outcome) and Object Type (truck or ball as Target Object) as between-subjects factors revealed a significant main effect of Condition on infants’ exploration preferences, $F(1, 42) = 5.23, p = .03, \eta^2 = .11$, no main effect of Object Type, and no interaction between Condition and Object Type. Infants who had seen the Surprising outcome of the solidity event preferred to explore the familiar Target Object ($M_{\text{preference}} = 3.76, SEM = 1.14$), a preference that differed from chance, $t(23) = 3.30, p < .01$. In contrast, infants who saw the Expected outcome showed no such preference ($M_{\text{preference}} = .09, SEM = 1.19$), $t(21) = .08, p = .94$ (Fig. 3).

2.3. Experiment 1 Discussion

The results of Experiment 1 suggest that surprise motivated infants to explore. Infants who saw an object violate solidity explored that object significantly more than infants who saw the same object behave in an expected way. This replicates the results of previous studies (Sim & Xu, 2017; Stahl & Feigenson, 2015), and suggests that violations-of-expectation are not only detected by infants, but also change their information-seeking behavior.

However, it remains unknown why surprise triggers infants’ exploration. It may be that surprise-induced exploration merely reflects arousal, straightforwardly increasing infants’ engagement with surprising objects. Alternatively, infants’ exploration might reflect the search for an explanation for the surprising event. If explanation-seeking is what motivates infants to explore, then providing infants with an explanation following a surprising event might be expected to abolish surprise-induced exploration. Previous research finds a related pattern of behavior in preschool-aged children, who stop asking causal questions once they have received an explanation, but persist following non-explanations (Frazier et al., 2009).

To find out whether preverbal infants recognize explanations, in Experiment 2 we showed a new group of infants the surprising solidity event from Experiment 1. As before, infants saw a moving object apparently pass through a red wall and come to rest on the wall’s far side. Then, as before, the experimenter rotated the wall so that infants could see its front face. However, this time, rotating the wall revealed that it contained an opening large enough for the object to have passed through — thereby providing a plausible explanation for the surprising event (i.e., the object passed through the hole). If infants’ surprise-induced exploration reflects the search for an explanation, then seeing the wall with the opening in it should decrease infants’ subsequent exploration, relative to that exhibited by infants in Experiment 1 who received no such explanation.

3. Experiment 2

3.1. Method

3.1.1. Participants

Twenty 11-month-old infants participated (mean age: 11 months, 9 days; 8 girls; range: 11 months, 5 days to 13 months, 24 days). Data from seven additional infants were excluded for fussiness (2), parent interference (1), failure to look at the event outcome (3), or failure to engage

---

1. Planned analyses of visual and manual exploration, coded separately, are reported in the Supplementary Materials.
in any object exploration (1).

3.1.2. Procedure
Infants first saw a Familiarization trial as in Experiment 1. All infants then saw the same solidity event as infants in the Surprising condition of Experiment 1: the Target Object was placed at the top of the ramp and then released so that it rolled down and passed behind the occluding screen, above which the red wall protruded. The experimenter then lifted the screen to reveal that the object had come to rest on the far side of the wall, as though it had passed through it. As in Experiment 1, infants had 10 seconds to view this surprising outcome. Then the experimenter reached down and picked up and rotated the red wall, placing it so that its front face was visible to infants. Whereas in Experiment 1 the red wall was revealed to be solid (Fig. 2, Panel 4A), in Experiment 2 the red wall was revealed to have a large opening in its face (Fig. 2, Panel 4B). Infants’ looking at the object and the wall was measured during the 20 seconds that followed, as described in Experiment 1.

The exploration phase that followed was as in Experiment 1. The experimenter emerged holding the Target Object (e.g., truck) and a novel Distractor Object (e.g., ball), said, “Look at these!” and placed both objects on the high-chair tray. Infants had 30 seconds to explore the objects.

3.2. Results
We first asked whether infants in Experiment 2 detected the solidity violation by comparing their looking at the event outcome to that of the infants in Experiment 1 who had seen an Expected event. We found that infants in Experiment 2 looked at the target object (e.g., the truck) for 5.37 s (SD = 2.40); this was significantly less than the 6.77 s (SD = 1.46 s) looking at the Target Object by infants in the Expected Outcome condition of Experiment 1 (M_{difference} = 1.54, SEM = .49, t(40) = 3.12, p < .01, 95% CI [.54, 2.54]). In contrast, there was no significant difference in looking between infants in Experiment 2 and infants in the Surprising Outcome condition of Experiment 1 (M_{difference} = .39, SEM = .64, t (42) = .63, p = .53, 95% CI [−.88, 1.68]). Hence infants in Experiment 2 detected the solidity violation.

Next we examined infants’ exploration, first computing exploration preference scores (i.e., Target Object exploration minus Distractor Object exploration). Because we hypothesized that infants’ exploration following a surprising event would depend on whether they had seen a plausible explanation for the event, we compared the exploration preference scores of infants in Experiment 2 with those of infants in Experiment 1. A 2 × 2 ANOVA with Condition (Surprising Outcome in Experiment 1 vs. Surprising Outcome with Explanation in Experiment 2) and Object Type (truck or ball as Target Object) as between-subjects factors revealed a main effect of Condition on infants’ exploration preferences, (F(1, 40) = 6.23, p = .02, \eta^2 = .13), as well as a main effect of Object Type (F(1, 40) = 4.72, p = .04, \eta^2 = .11), and no interaction between these. Infants in Experiment 2, who had received an explanation for the surprising event, explored the target object significantly less (M_{preference} = −1.62, SEM = 1.59) than infants in the Surprising Outcome condition of Experiment 1, who received no explanation, (M_{difference} = −5.37, SEM = 2.15, 95% CI [−9.73, −1.02]).

Did seeing a plausible explanation make infants treat a surprising event like an expected event? We next compared the exploration preferences of infants in Experiment 2 to those of infants in the Expected Outcome condition of Experiment 1. A 2 × 2 ANOVA with Condition (Surprising Outcome with Explanation in Experiment 2 vs. Expected Outcome in Experiment 1) and Object Type (truck or ball as Target Object) as between-subjects factors found no main effect of Condition on infants’ exploration preferences, F (1, 38) = .50, p = .48, \eta^2 = .01, no main effect of Object Type, and no interaction between Condition and Object Type. As predicted, infants explored equally after receiving an explanation for a surprising outcome and after seeing an expected outcome (M_{difference} = −1.60, SEM = 2.27, 95% CI [−6.20, 2.99]; Fig. 3).

Did seeing the wall with the hole in it change infants’ looking simply because it was a perceptually novel object (perhaps more novel than the solid wall in Experiment 1), rather than by virtue of the potential explanation it provided? If infants in Experiment 2 found the wall especially interesting, they might have spent more time looking at the stage area during the 20 seconds after the rotated wall was revealed. If this looking included not just the wall but also the Target object itself, then perhaps infants in Experiment 2 simply received more exposure to the Target object than infants in Experiment 1, and were subsequently less interested in exploring it. Infants’ looking patterns argue against such an interpretation. During the 20 seconds when both the rotated wall and the object were visible, infants in the Surprising Outcome condition of Experiment 1 looked at the target object (e.g., the truck) for 3.99 s (SD = 2.64), and infants in the Surprising Outcome with Explanation condition of Experiment 2 looked at it for 4.16 s (SD = 2.40; this

---

**Fig. 2.** Schematic of Experiment 2. The experimenter released the Target Object at the top of the ramp (Panel 1). The Target Object rolled down the ramp and disappeared behind an occluding screen, above which a bright red wall protruded (Panel 2). The screen was then removed, revealing the Target Object resting on the far side of the red wall (Panel 3). After 10 s the experimenter rotated the red wall. Whereas infants in the Surprising condition of Experiment 1 saw that it was solid (Panel 4A), infants in Experiment 2 saw that the red wall contained a large opening (Panel 4B).
looking difference was not significant, $t(42) = .21, p = .83$. Infants looking at the wall also did not differ across experiments: infants in the Surprising Outcome condition of Experiment 1 looked at the solid wall for 4.03 s ($SD = 1.75$), and infants in the Surprising Outcome with Explanation condition of Experiment 2 looked at the wall with the hole in it for 4.79 s ($SD = 2.09$); this difference was not significant, $t(42) = 1.33, p = .19$.

4. Experiment 3

The main finding from Experiment 2 was that seeing an explanation for a violation-of-expectation event eliminated infants’ exploration of the object that had previously behaved surprisingly. In Experiment 3, we sought to replicate this finding with a new sample of infants, to examine the robustness of this effect.

4.1. Method

4.1.1. Participants

Twenty 11-month-old infants participated (mean age: 11 months, 4 days; 7 girls; range: 10 months, 3 days to 12 months, 26 days). Data from two additional infants were excluded for fussiness (1) and parent interference (1).

4.1.2. Procedure

The procedure was exactly as in Experiment 2.

4.2. Results

First we asked whether infants in Experiment 3 detected the solidity violation by comparing their looking at the event outcome to that of infants in Experiment 1. Infants in Experiment 3 looked significantly longer at the event outcome ($M = 6.57$ seconds, $SD = 1.56$ seconds) than infants in the Expected condition of Experiment 1 ($M_{\text{difference}} = 1.34$ seconds, $SEM = .51, t(40) = 2.65, p = .01, CI [.32, 2.37]$). There was no difference in looking at the event outcome between infants in Experiment 3 and infants in the Surprising condition of Experiment 1, ($M_{\text{difference}} = -.59$ seconds, $SEM = .64, t(42) = .92, p = .36, CI [-.190, .719]$), affirming that infants in Experiment 3 detected the solidity violation.

Next we analyzed infants’ exploration preference scores. A $2 \times 2$ ANOVA examined the effects of Condition (Surprising Outcome in Experiment 1 vs. Surprising Outcome with Explanation in Experiment 3) and Object Type (truck or ball as Target Object) on infants’ exploration. This revealed a marginally significant main effect of Condition, $F(1, 40) = 3.95, p = .05, \eta^2 = .09$, no effect of Object Type, and no interaction between Condition and Object Type. Infants explored marginally less in Experiment 3 after seeing a surprising outcome accompanied by an explanation ($M_{\text{preference}} = -.12, SEM = 1.44$) than in Experiment 1 after seeing the same surprising outcome but without any explanation ($M_{\text{difference}} = 1.95, 95\% CI [-.783, .666]$, Fig. 3). As with Experiment 2, this shift in preference was not driven by differences in infants’ looking immediately following the wall rotation, compared to infants in Experiment 1. During the 20 seconds when both the rotated wall and the object were visible, infants in the Surprising Outcome with Explanation condition of Experiment 3 looked at the target object for 4.09 s ($SD = 2.40$) and the wall with the hole in it for 4.22 s ($SD = 1.75$). Infants in the Surprising Outcome condition of Experiment 1 looked at the target object for 3.99 s ($SD = 2.64$) and the solid wall for 4.03 s ($SD = 1.75$). Neither of these cross-experiment comparisons was significant (differences in infants’ looking at the target object, $t(42) = .13, p = .90$; difference in infants’ looking at the wall, $t(42) = .07, p = .94$).

To ask whether infants treated a surprising event accompanied by an explanation as an expected event, we compared infants’ exploration in Experiment 3 to infants’ exploration in the Expected Outcome condition of Experiment 1. A $2 \times 2$ ANOVA with Condition (Expected Outcome in Experiment 1, Surprising Outcome with Explanation in Experiment 3) and Object Type (truck or ball as Target Object) as between-subjects factors revealed no effect of Condition, $F(1, 38) = .00, p = .96, \eta^2 = .00$, no effect of Object Type, and no interaction between these. Thus, like infants in Experiment 2, infants who saw a surprising event followed by a plausible explanation for that event showed exploration preferences that were no different from infants who had never seen a surprising event in the first place.

Finally, as a further test of whether explanations drove infants’ exploratory behavior, we examined individual differences in infants’ looking times and their possible link to subsequent exploration. Because we had hypothesized that seeing an explanation for a surprising event would abolish surprise-induced exploration, we also predicted that the longer infants looked at the explanation (i.e., the red wall with the opening in it), the less they would subsequently explore the Target Object. That is, although infants in Experiments 2 and 3 did not look longer at the wall overall than infants in Experiment 1, we predicted that infants who spent more time processing the explanation might be less driven to explore the object that had behaved surprisingly. We tested this prediction using linear regression. First, because infants saw identical sequences of events across Experiments 2 and 3, we combined these
two samples to empower a more robust analysis of potential individual differences. Before we did so we confirmed that group performance did not differ between the two experiments. Infants in Experiments 2 and 3 looked equally long the initial event outcome ($M_{\text{difference}} = -.20, SE = .48, t (38) = .41, p = .68, 95\%\ CI [-1.17, .77]$). Similarly, a $2 \times 2$ ANOVA with Condition (Surprising Outcome with Explanation in Experiment 2 vs. Surprising Outcome with Explanation in Experiment 3) and Object Type (truck or ball as Target Object) as between-subjects factors found no main effect of Condition on infants’ exploration preferences, $F (1, 36) = .33, p = .57, \eta^2 = .01$, no main effect of Object Type, and no Condition by Object Type interaction. Given this lack of differences between Experiments 2 and 3, we collapsed the two to create a larger sample of 40 infants (but see Supplementary Materials for separate analyses by experiment). Among this collapsed group, we found that the longer infants looked at the explanation (the rotated wall with the hole in it), the more they preferred to explore the novel Distractor Object, $R^2 = .13$, $F (1, 36) = 5.59, p = .02, B = -1.46 (SE = .62), 95\%\ CI [-2.72, -2.11]$ (Fig. 4A). No such relationship was found among infants in either condition of Experiment 1, for whom the red wall offered no explanatory information (Surprising condition of Experiment 1, $R^2 = .02$, $F (1, 22) = 4.5, p = .51, B = .45 (SE = .67), 95\%\ CI [-.94, 1.84]$) (Fig. 4B): Expected condition of Experiment 1, ($R^2 = .02, F (1, 20) = .46, p = .51, B = -.48 (SE = .71), 95\%\ CI [-.96, 1.00]$).

5. General discussion

Here we find that long before they can ask or answer “why” questions, infants seek and recognize explanations for anomalous events. In our first experiment, we confirmed that 11-month old infants who saw an object violate a principle of typical object behavior (by passing through a solid wall) significantly increased their tendency to explore that object, relative to infants who saw a closely matched event in which the object accorded with expectations (replicating previous work; Stahl & Feigenson, 2015). Then, in Experiment 2, we showed that this surprise-induced exploration reflected infants’ search for an explanation: when provided with a plausible reason for the surprising event, infants’ exploratory preference disappeared, causing them to treat the previously surprising event like an expected one. We replicated these findings in Experiment 3. Our results suggest that infants’ surprise-induced exploratory activity is not triggered ballistically by violation events, and does not merely reflect heightened arousal. Instead, infants appear to be actively searching for explanatory information, such that when that information is found, exploration returns to baseline.

We also saw that infants’ ability to recognize explanations for objects’ surprising behavior differed among individuals. As a group, infants in Experiments 2 and 3 who had seen an explanation for the surprising event showed no significant exploratory preference—they spent roughly equal amounts of time playing with the object from the surprising event and the novel distractor. However, closer inspection revealed that some of these infants in fact had a strong preference to explore the target object, whereas others had a strong preference to explore the distractor. Critically, which preference infants exhibited was predicted by how long they looked at the explanation for the surprising event: infants who looked at the explanation longest showed the greatest tendency to ignore the previously surprising object and explore the distractor (that is, showed a reverse preference relative to infants in Experiment 1 who received no explanation). The source of these individual differences remains an important direction for future study. An intriguing possibility is that infants differed in their ability to recognize the wall with the hole in it as a satisfying explanation for the surprising event, and/or in their desire to find an explanation for an expectancy violation in the first place. While previous studies have found stable individual differences in the rate of infants’ habituation and in their interest in perceptual novelty (Bornstein & Sigman, 1986; Colombo, Mitchell, Coldren, & Freeseman, 1991; Fagan, 1984), little is known about the possibility of individual differences in infants’ ability to detect violations of expectation (but see Perez & Feigenson, 2019a, 2019b) or in their ability to recognize explanations for such violations (as in the present work). Our findings therefore highlight the study of individual differences in response to violations, and their potential impact on later cognition, as an important avenue for future research. To support a robust analysis of individual differences among infants, such future

![Fig. 4. A. Relationship between infants’ looking at an explanation following a surprising solidity event and their exploration preferences in Experiments 2 and 3. B. Relationship between infants’ looking at a non-explanation following a surprising solidity event and their exploration preferences in Experiment 1.](image-url)
work will certainly want to examine samples larger than those reported here.

Although we argue that infants in Experiments 2 and 3 recognized the wall with the opening in it as an explanation for the apparent solidity violation, there is another, deflationary interpretation of our findings: that infants who saw this wall were distracted by its unusual shape, and this distraction led them to forget about the surprising event just prior, thereby causing their lack of exploration preference when given the Target and Distractor objects to explore. There are several reasons we think the wall with the hole in it served as an explanation, rather than a distraction. First, infants who looked longest at the wall with the hole in it were also the infants who were most likely to ignore the Target Object. Yet infants in the Surprising Outcome condition of Experiment 1 did not show this effect: infants who looked longest at the solid wall were not the most likely to later ignore the Target Object. This argues against interest in the wall (independent of its ability to explain the solidity violation) predicting infants’ subsequent exploration. Second, although it could potentially be argued that only the wall with the hole in it distracted infants from processing the solidity violation (and that although infants looked equally at the solid wall, it diverted their processing to a lesser degree), several previous studies find that infants are in fact able to reason about entities similar to our red wall, demonstrating an understanding that objects can pass through tunnels and barriers with openings (e.g., Côbra & Johnson, 2013; Schweinle & Wilcox, 2004). These results suggest that infants are not merely confused by these entities. Rather, they have expectations about how an object’s shape influences its interactions with other objects. Still, future work might implement tighter controls to definitively rule out this interpretation. One such design would simply flip the wall with the opening upside-down, so that the opening is at the top rather than the bottom. We predict that even though this U-shaped wall is just as novel as the inverted U we used in Experiments 2 and 3, infants would reject this as an explanation for a solidity violation, and show a surprise-induced tendency to explore the Target Object.

If infants in our study were in fact seeking and recognizing explanations, how is this ability related to the kinds of explanations that have been studied in older children and adults? Violations and explanations that can be represented using the core knowledge systems that are available to preverbal infants and nonverbal animals (Carey, 2007; Spelke & Kinzler, 2007) may or may not be linked to the verbally articulated explanations sought by older children, often to account for more complex phenomena and involving explanations that would be harder to uncover through a process of individual discovery (e.g., “Why is the sky blue?” “How do bees grow their babies?”) (Chouinard, Harris, & Matsos, 2007). One approach to begin characterizing the relationship between explanation-seeking in infancy and explanation-seeking in later childhood is to examine individual differences over time—for example, asking whether infants who are most interested in explanations for surprising events later become children who search most persistently for verbal explanations. If these are linked, this could implicate the existence of a stable trait—one potentially related to the constructs of curiosity or “desire for understanding” that have been explored in older children and adults (e.g., Berlyne, 1966; Litman, Hutchins, & Russon, 2005; Loewenstein, 1994; Mussel, 2010), and which might even be measurable early in life.

The present research also raises the question of whether infants represent some explanations as better than others. To illustrate, imagine four different explanations for the solidity violation in the present study: one where the red wall is revealed to contain a large opening that is plenty wide enough for the object to have passed through, another where the red wall contains an opening barely wider than the object itself, a third where the red wall contains an opening considerably narrower than the object, and finally one where the red wall contains a wide opening but on its top surface (as in a U) rather than its base. These scenarios vary in the likelihood that they can provide an adequate explanation for the surprising event of the ball passing through the wall. Would infants’ interest in these explanations, and their change in subsequent exploration, decrease as a function of the explanations’ goodness? Adults are sensitive to such variations in “explanatory power,” where information is said to have more explanatory power as it renders observations less and less surprising (Colombo, Bucher, & Sprenger, 2017; Schupbach & Sprenger, 2011). Older children too are sensitive to explanatory power, in that they recognize explanations that, because of their circularity, are wholly unsatisfying (Baum, Danovitch, & Keil, 2008; Castelain, Bernard, & Mercier, 2018). A small amount of evidence suggests that infants’ responses to expectancy violations are not all-or-none; rather, their looking increases with the surprisingness of the stimulus event (Téglás et al., 2011). As pointed out by Sim and Xu (2019), these modulations in infants’ responses may reflect the extent to which infants experience evidence for a model other than the one they previously held (or, as discussed by Foster and Keane (2015) for the case of adults, the extent to which observations are difficult to explain on an existing model). This could also be true of infants’ interest in potential explanations of surprising events: explanations that best support a clear alternative hypothesis may exert the strongest effect on infants’ subsequent behavior. However, infants’ ability to represent degrees of explanation in this way remains to be explored.

Our results also highlight a question that will be vital for future work: what aspects of perceptual and cognitive processing are revealed by infants’ longer looking? In the current experiments, we found that infants looked longer at an event that culminated in a surprising compared to an expected outcome—a pattern seen across hundreds of studies using violation-of-expectation. This elevated looking at events that adults describe as anomalous (or even impossible) has sometimes been described in terms of infants experiencing “surprise,” or “puzzlement,” or as reflecting a prediction error. And yet, we also observed that, for infants in Experiments 2 and 3, longer looking at an explanation for a surprising event was correlated with infants treating the event as unsurprising. The longer infants looked at a wall with a hole in it, having just seen an object appear to pass through, the less they explored that object when given the opportunity. In this case, longer looking seems to have reflected something other than surprise or puzzlement—for example, deeper processing of the new information, such that surprise was actually diminished. The idea that longer looking reflects more than a single underlying response or process has been discussed in other work (Aslin, 2007; Powell & Saxe, forthcoming) and in principle applies to any measure, from behavioral to neural. When distinct processes are reflected in a single measure, these can sometimes be disentangled by asking whether separate brain areas are active during the time the measure is being obtained (Powell & Saxe, forthcoming), or by finer grained temporal analyses. In general, multiple measures may be useful, or required (LoBue et al., 2020), for understanding the relationship between infants’ responses to predicted events, to surprising violations of these predictions, and to explanations of the violations.

In summary, our findings demonstrate that infants do not merely detect surprising events, they also use surprise as inspiration to seek explanations (see also the discussion of Explanation-based Learning by Baillargeon & DeJong, 2017). Starting in the first year of life, young learners leverage differences between what they predict and what they observe to motivate the gathering of new evidence and the revision of their beliefs. Whereas the ability to test explanations in a systematic way, and to generate novel explanations, likely undergoes considerable development during later childhood, the basic capacity to search for and recognize explanations is functional early in life.

Funding resource

JP was supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. (DGE-1746891). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.
CRediT authorship contribution statement

Jasmin Perez: Conceptualization, Methodology, Formal Analysis, Investigation, Visualization, Writing-original draft, Writing-review & editing. Lisa Feigenson: Conceptualization, Methodology, Visualization, Writing-review & editing.

Acknowledgements

We thank the parents and children who participated in this study, as well as Madeleine Tan Cuan for their help in data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2021.104942.

References