Published in final edited form as:
Infancy. 2013 ; 18(6): . doi:10.1111/infa.12014.

Expectations about single event probabilities in the first year of life: The influence of perceptual and statistical information

Chris A. Lawson* and
University of Wisconsin-Milwaukee

David H. Rakison
Carnegie Mellon University

Abstract

Recent evidence suggests that infants can generate expectations about future events from a sample of probabilistic data. However, little is known about the conditions that support the development of this ability. Three experiments tested the prediction that 8- and 12-month-olds respond to base rates as well as perceptual cues when they generate expectations from a sample of probabilistic data. Results revealed that 12-month-olds were sensitive to the statistical and perceptual properties of the evidence depending on the distribution of high-to-low base rate items in the sample. Specifically, 12-month-olds focused on perceptual features of the evidence when a sample was large and more skewed (e.g., 6:1), whereas they attended to statistical properties when the sample was smaller and less skewed (e.g., 4:1). In contrast, eight-month-olds always focused on the perceptual features of the evidence. Neither group generated expectations from a small, less skewed sample (e.g., 2:1). These results suggest that the ability to generate expectations about future events is mediated by specific features of the available evidence and undergoes significant change during the 1st year of life.

Keywords

probabilistic inference; infants; future events; probability judgment

---

The ability to generate expectations about future events from probabilistic evidence is critical to any rational decision-making process (Anderson, 1991; Tenenbaum & Griffiths, 2001). Recent findings suggest that by the end of the first year of life infants have acquired some of the basic skills at the core of rational inference: Prior to their first birthday infants are able to generate the expectation that a sample will yield the most likely event given the distribution of evidence in the sample (Denison & Xu, 2010a; Teglas, Girotto, Gonzalez, & Bonatti, 2007; Teglas, Vul, Girotto, Gonzalez, Tenenbaum, & Bonatti, 2011; Xu & Garcia, 2008). Only a small number of these studies have explored the factors that support infants' ability to generate expectations about future events or the developmental progression of this skill. To address these issues, the present studies tested the possibility that base rate information and the available perceptual cues influence the ability of 8- and 12-month-old infants to form expectations about the outcomes yielded from samples of probabilistic data.

Teglas et al. (2007) conducted one of the first studies on infants' expectations about single events drawn from a sample of probabilistic evidence. When presented a container with 3 identical objects and 1 object of a different color and shape 12-month-olds looked longer

---

*corresponding author: Chris Lawson; University of Wisconsin-Milwaukee; Department of Educational Psychology; Enderis Hall; 2400 E. Hartford Ave., Milwaukee, WI 53211; lawson2@uwm.edu.

when the container yielded a low probability event (i.e., the object of a different color and shape exited the container) than when the container yielded a high probability event (i.e., one of the 3 identical objects exited the container). In a control condition in which a divider in the container made it physically impossible for one of the three similar objects to exit, infants did not look longer when the single object exited the container (see Denison & Xu, 2000b for similar results). The authors interpreted these findings to mean that 12-month-olds distinguish between possible and impossible events, and that they expect a high probability event is more likely than a low probability event (see also Teglas et al. (2011), and Denison & Xu (2010a) who showed similar results with 12-month-olds in a preferential choice task).

Other evidence suggests the ability to use probabilistic evidence to make inferences about future events emerges before 12 months. For example, in a study by Xu and Garcia (2008), 8-month-olds were familiarized to a large box that contained 70 balls of one color and 5 balls of a different color (e.g., 70 red balls, and 5 white balls). The contents of the box were then concealed and an experimenter " randomly" selected a sample of 5 balls from the population. On some test trials the sample contained a distribution of balls consistent with the distribution in the population (e.g., 4 red balls and 1 white ball) and on other trials the sample included a distribution that was inconsistent with the population (e.g., 4 white balls and 1 red ball). Eight-month-olds looked longer when the inconsistent sample was drawn from the population, presumably because this sample violated their expectation that the chosen sample would have the same distribution as the population from which it was selected (Denison, Reed, & Xu, in press presented similar results with 6-month-olds). In a control condition in which the sample was drawn from the pocket of the experimenter and then placed near the population of balls, infants showed no preference for either the consistent or inconsistent samples, indicating the mechanism by which infants match samples to populations is sensitive to the sampling procedure.

One conclusion from this work is that infants pay close attention to the base rate information – or distribution of high-to-low probability items – to generate expectations about the most likely future event. From this perspective infants should be more disposed to expect a high probability event when the distribution of low-to-high probability items is more skewed than when the distribution is less skewed. One of the goals of the present studies was to test this prediction by varying the distributions of low-to-high probability items within the samples presented to infants (e.g., 2 vs. 1, 4 vs. 1, and 6 vs. 1).

The current studies also explored the influence of perceptual cues, such as the location of the balls in the container, on expectations about future outcomes. It would be surprising if infants did not consider perceptual information when generating expectations about future outcomes given the evidence detailing the integral role of perceptual information in infants' expectations about events (e.g., Bremner, Slater, Mason, Spring, & Johnson, in press; Johnson, 2010; Kellman & Spelke, 1983). Indeed, Teglas et al. (2011) showed that under some conditions infants expectations about future outcomes are based on the location of balls in a sample rather than the available base rate information (see also Denison & Xu, 2000b). In their study, 12-month-olds were shown a container with 4 moving objects (3 high probability items and 1 low probability item) and then the containers were covered for .4 s, 1 s, or 2 s. After this occlusion period infants were shown an event in which the low probability ball exited the container and another in which the high probability ball exited of the container. In the .4 s and 1 s condition infants looked longer when the exiting ball was perceptually incongruent with the ball that was expected to be in the location nearest the exit; in other words, infants were surprised when the exiting ball was a different color from the ball located closest to the exit of the container. In the 2 s condition the infants looked longer when a high probability item exited the container, regardless of the location of balls in the container, presumably because ball location was unpredictable in this condition. These findings reveal two things about infants' expectations about probabilistic samples. The first is that by 12 months of age infants exhibit some cognitive flexibility in their ability to incorporate and use multiple cues from a sample of probabilistic evidence. The second is that whether infants focus on perceptual or statistical properties of the evidence is dictated by their expectations about moving objects. It is possible other cues influence whether infants are drawn to the perceptual, rather than base rate, information in the sample. The current studies explored the possibility that attention to ball location would be mediated by the distribution of low-to-high probability items in the sample.

Figure 1 provides a schematic overview of the three experiments reported in this paper. Several features of these experiments are noteworthy. First, the stimuli included two boxes each of which contained an inverse proportion of high-to-low base rates (see Figure 1 panel 6). Including an equal number of red and blue balls on the screen eliminated the possibility that looking behavior would be guided by the perceptual mismatch between the single test event and the objects on the screen (for a similar rationale see Denison et al., in press). Second, the test events were designed so that the balls exiting the containers represented (1) either low or high probability events relative to the contents of the container from which the balls exited, and (2) perceptually congruent or incongruent events relative to the ball within the container closest to the exiting ball. Third, the disparity between the high-to-low probability items within the containers was manipulated across the experiments to examine infants' sensitivity to different base rates in the available evidence (e.g., 2:1, 4:1, and 6:1). Fourth, the current experiments employed a visual habituation paradigm to ensure better encoding of the input and to be sure looking behavior was based on expectations about the novelty of the test events (Hunter & Ames, 1988). Finally, to understand the development of infants' expectations about probabilistic samples we included two age groups, 8- and 12-month-olds.

We predicted this experimental design would elicit one of three patterns of looking. The first is a probability-based prediction: Expectations based the distribution of properties should lead to longer looking times for low probability outcomes in comparison to high probability outcomes because the former would violate the expectation that the sample will yield the most likely event. Furthermore, if infants expect a high probability outcome based on the distribution of evidence in the samples, they may be more likely to do so for the most disparate sample (e.g., 6:1) than the less disparate sample (e.g., 2:1), because the former skews the distribution of items in the sample and thus makes the high probability outcomes more likely.

The second is a perceptual-based prediction: Expectations based on the location of balls in the container should lead to longer looking times for events in which the exiting ball is perceptually incongruent with the adjacent ball because this event violates their expectation that the item closest to the exit will emerge from the container.

The third, mixed, pattern was expected to emerge assuming that some conditions elicit attention to the location of balls in the containers while others elicit attention to the distribution of balls in the containers. One reason to expect a mixed pattern is because the high disparity conditions include a large number of items on the screen (e.g., the 6:1 condition included 14 items on the screen) and thus may exceed the processing capacity of both groups of infants. For this reason, because the items located nearest the exit of the container appear early during the habituation trials (see Figure 1, panel 4) infants are more likely to encode this feature of the input. The greater disparity within these large samples may also enhance the perceptual contrast between low and high probability items and thus may cause infants to focus on ball location. Thus, the high disparity condition may be more likely to elicit a perceptual-based looking pattern, while the less disparate samples may not.

One implication from this view is that smaller, less disparate samples are more likely to yield the probability-based pattern because the infants will be able to encode all of the items in the samples and because the perceptual contrast between items is not as strong. However, a smaller sample does not guarantee probability-based expectations. There is less disparity between high and low probability items within the smaller sample (e.g., Experiment 2 included a sample of 2:1) and thus the low probability event is less likely relative to the low probability event for a more disparate sample (e.g., Experiment 1 contains a sample of 4:1). For this reason less disparate samples may be more likely to yield random looking patterns.

In short, the present studies were designed to test these different predictions about the influence of base rate information on infants' use of perceptual and statistical properties of evidence to generate expectations about future outcomes.

Experiment 1

In this experiment, 8- and 12-month-old infants were habituated to events during which two containers were filled with inverse ratios of blue and red balls (e.g., 4:1 and 1:4). Following habituation infants were shown a low and a high probability event in which a single ball exited each of the now opaque containers. Additionally, the test events varied whether the exiting ball and the ball most adjacent to the ramp were either the same color (e.g., both blue) or different colors (e.g., one was blue while the other was red) (see Figure 1). We predicted that if infants are able to generate the expectation that each container will yield a high probability event they should look longer at the low probability events because such an event would violate their expectation. In contrast, if infants focus on ball location rather than the base rates in each box, they should look longer at events in which the exiting balls is a different color than the ball nearest the exit of the container, because this event would violate their expectation that the ball closest to the exit would emerge from the container.

Method

Participants—Participants were 16 healthy full-term 8-month-olds (*M*=8 months, 21 days; range=8:1 to 9:14; 8 males, 8 females) and 16 healthy full-term 12 month-olds (*M*=12 months, 7 days; range=11: 21 to 13:12; 8 males, 8 females). Data from an additional 9 infants were excluded from the final sample due to failure to habituate (1), fussiness (3), and failure to meet criterion for inclusion in the data analysis (5) (see below for a description of inclusion criterion). Infants were recruited through birth lists acquired from a private company. Families were given a small gift for their participation. In this experiment, and the others reported here, the majority of infants were White and of middle socioeconomic status.

Design and stimuli—The habituation and test stimuli were computer-animated events created with Macromedia Director 8.0. Each event included a scene with two boxes; one located approximately 10 cm from the bottom left edge of the computer screen and the other approximately 10 cm from the bottom right edge of the computer screen. Individual balls entered from the top of the screen and dropped simultaneously into each of the boxes at an interval of one ball per second. The dropped balls remained visible in the boxes. A total of five balls landed in each of the boxes. One box contained 4 blue balls and 1 red ball, while the other box contained 4 red balls and 1 blue ball. Thus, an equal number of red balls and blue balls appeared on the screen (i.e., 5:5). After the last balls landed the contents of the boxes remained visible for 1 second and then a screen covered the contents of each box. The covered boxes remained on the screen for 2 seconds. Each event lasted 10 seconds and was repeated up to three times within a single video. The habituation phase included 16 videos in which the low probability ball was the first, second, third, or fourth ball dropped into the containers. The low probability ball was never the final ball dropped because such an event would involve a scenario in which the ball would appear to go through the ball placed on the top row of the container. These 16 videos were counterbalanced for each participant to increase the likelihood each participant would be habituated to the same events.

For the test trials the boxes were full when trial was initiated and the distribution of balls in each box was left visible for 3 seconds after which a screen came up to cover the contents of each box (see Figure 1, panels 7 and 8). After a 1 second occlusion period a test event was presented. Each infant was presented four test trials. These test trials crossed probability of outcome and perceptual congruity between the exiting object and the location of balls in the container. In the *Low probability* trials the ball that occurred with the lowest probability within a box exited each of the boxes and in the *High probability* trials the ball that occurred with the highest probability within each box exited each of the boxes. In *congruent* trials, the exiting ball was the same color as the most adjacent ball (see Figure 1, frame 10), whereas in *incongruent* trials, the exiting ball was a different color from the most adjacent ball (see Figure 1, frame 11). Thus each infant was presented the following four test trials: *low probability/congruent, low probability/incongruent, high probability/congruent*, and *high probability/incongruent*. Test events were counterbalanced across participants so that each event was presented an equal number of times as the first, second, third, and fourth event.

Because the test phase included a period of 3 seconds in which children were shown the contents of the box and 1 second during which the balls were occluded, we excluded from data analyses infants whose looking time during the test phase failed to exceed 4 seconds for at least one of the test trials (N=5).

Apparatus—Each infant was tested in a dimly lit laboratory room. During the procedure trials appeared on a 43-cm television monitor approximately 80 cm from the infant’s face. Behind the monitor and surrounding the testing area was a black curtain that reached from the ceiling to the floor. A video camera was located behind the curtain and was concealed from view by the black curtain. The video camera was connected to a television monitor that displayed infants’ visual fixation. The monitor was connected to a video recorder allowing infants’ looking behavior to be observed by the experimenter.

Procedure—We used subject-controlled criterion to establish habituation. The criterion required that the mean looking time for 3 consecutive trials was less than 50% of the mean looking time for the first 3 trials. Looking behavior was recorded by keystroke by a trained researcher who observed the experimental trials from a computer connected to a video camera fixated on the infant. The trials were initiated when an infant looked at the monitor. When an infant looked away from the monitor for more than 1 s an attention getter was initiated to redirect their attention to the monitor. A trial was terminated if an infant looked away from the screen for 1 s or if an infant continued to look at the screen for the entire 30 s for a given trial. A new trial was initiated when attention was drawn back to the monitor. The attention getter was a video with an expanding green ball accompanied by a loud beeping sound. The experimental software administered trials based on the established criterion and recorded the looking time data.

Results and Discussion

With the exception of one infant who was not included in the final analyses, all participants reached the established habituation criterion. On average, 8-month-olds reached habituation in 97.23 s (SD=64.3 s) and were exposed to approximately 7 events (M=7.19, SD=2.81) whereas 12-month-olds habituated on average in 108.57 s (SD=83.71 s) and were exposed to approximately 8 events (M=8.20, SD=3.71). Separate t-tests revealed no significant differences between 8- and 12-month-olds in the duration of the habituation trials t(15)=.76,

p=.46, or in the number of videos observed, t(15)=.44, p=.67, which indicates that both groups of infants were exposed to the same amount of input during habituation.

The next analysis examined potential order effects of the test events. Recall that the habituation videos did not include events in which the ball exited the container. Unless infants have a-priori expectations about the type of events afforded by the containers they must learn how these containers work during the experiment. For this reason we expected the duration of looking times would decrease over the series of test trials, such that infants would look longer at the first trial because the exiting ball is a novel event, while looking time would be shorter for later trials for which the event is no longer novel. In support of this prediction, a mixed-design ANOVA with Age (8-month-olds, 12-month-olds) as a between subjects variable and Test Order (first, second, third, fourth) as a within subjects variable revealed a main effect of Test Order, $F(3,90)=23.65$, $p<.001$, $\eta^2=.44$. Both age groups showed a decrease in looking times for each subsequent test event, though the only significant contrasts were between looking times for the first test event (M=10.06 s, SD=5.65 s) and each of the other three test events; second (M=6.44 s, SD=4.46 s), third (M=5.78 s, SD=4.43 s), and fourth (M=5.14 s, SD=3.89 s), all Fs>3.60, $p<.01$. Thus, infants in both groups recognized the novel exiting event during the experiment and did so relatively early in the testing phase.

To explore our primary hypotheses, mean looking times to the test trials were submitted to a mixed-design ANOVA with Age as a between subjects variable and Probability (High, Low) and Ball location (Congruent, Incongruent) as within subjects variables. The analysis revealed an effect of Ball location, $F(1,30)=18.40$, $p<.001$, $\eta^2=.38$, due to longer looking times to the Incongruent events (M=9.36 s, SD=2.92 s) than the Congruent events (M=6.18 s, SD=2.24 s). Supplemental analyses showed this effect of Ball location in younger infants, $F(1,15)=351.56$, $p< .001$, $\eta^2=.71$, but not in twelve-month-olds, $F<.45$, $p>.51$ (see Figure 2). These age differences were reflected in a Ball location by Age interaction, $F(1,30)=10.67$, $p=.003$, $\eta^2=.26$, with simple effects analyses indicating that older infants looked longer at Congruent events (than younger infants, $F(1,31)=9.69$, $p=.004$. Thus the data reveal a more robust effect of the perceptual congruence on the expectations of 8-month-old infants than on the expectations of 12-month-old infants.

The main analyses also yielded a Probability by Age interaction, $F(1,30)=4.56$, $p=.04$, $\eta^2=.13$. Twelve-month-olds looked at Low probability events significantly longer than 8-month-olds, $F(1,31)=4.26$, $p<.05$. There was no age difference in looking times for the High probability events. Additional analyses revealed that 12-month-olds looked at the Low probability events longer than the High probability events, $F(1,15)=6.22$, $p=.03$, $\eta^2=.29$ (see Figure 3). For 8-month-olds looking times for the Low probability events and High probability events were not significantly different, $F<1.30$, $p>.19$ (see Figure 3). These results indicate that the statistical information in the containers had a greater influence on the expectations of 12-month-olds than it did on the expectations of 8-month-olds.

Overall these results indicate that each group of infants relied on different sources of information to generate expectations about future events. Twelve-month-olds relied on base rate information rather than the location of balls in the container, whereas 8-month-olds relied on the location of balls in the container rather than the base rate information. One interpretation of these results is that the ability to utilize statistical information to generate expectations about future events develops sometime between 8 to 12 months of age. However, the results from Experiment 1 also support the interpretation that certain features of the available evidence, such as the base rates within the samples, influence whether infants attend to the statistical properties of the evidence or the perceptual properties of the evidence. The remaining experiments were designed to explore these two interpretations.

Experiment 2

It could be argued that 8-month-olds in Experiment 1 did not exhibit the probability-based looking pattern because their information-processing abilities were overloaded making it difficult for them to track and encode ten distinct objects represented in two different locations. Moreover, there may have been too many items within a single sample (n=5) for infants to be able to discriminate the low probability from the high probability instances. Evidence in support of this latter possibility comes from research that showed that 6-month-olds will discriminate dot arrays containing 3 items from those containing 2 items but cannot discriminate larger dot arrays, such as those that contrast 4 items versus 6 items (Starkey & Cooper, 1980; Xu, 2003; cf. Clearfield & Mix, 1999). Thus, 8-month-old infants in Experiment 1 may not have generated expectations on the basis of the probabilistic data because they were unable to discriminate 4 balls versus 1 ball in each of the boxes. Experiment 2 explored this possibility by replicating Experiment 1 with the exception that the habituation and test trials involved boxes that each contained 3 balls, each with an inverse ratio of 2:1. Thus, Experiment 2 also tested the extent to which 12-month-olds' expectations are sensitive to base rates. If expectations about future events are contingent on the magnitude of difference between high and low probability instances, then 12-month-olds may not generate expectations about future events from this less disparate distribution.

Method

Participants—The participants were 16 healthy full-term 8-month-olds (*M*=8 months, 11 days; range=7:22 to 9:12; 8 males, 8 females) and 16 healthy full-term 12 month-olds (*M*=12 months, 11 days; range=11:24 to 13:4; 8 males, 8 females). Data from an additional 10 infants were excluded due to fussiness (3), experimental error (2), and for failing to meet the criterion for inclusion in the data analysis (5). Methods of recruitment and reimbursement were the same as in Experiment 1.

Stimuli, design, apparatus, and procedure—The habituation and test trials involved events with three, rather than five balls. For the habituation trials the duration of ball drops was the same as Experiment 1 (5 s). However, the time was divided equally such that each ball drop lasted approximately 1.33 seconds. In all other respects the experiment was identical to Experiment 1.

Results and Discussion

Before conducting the main analysis it is informative to look at the duration of the habituation trials and potential order effects of the test events. On average, 8-month-olds reached habituation in 71.11 s (*SD*=14.49 s) and viewed approximately 7 habituation events (*M*=7.32, *SD*=3.01) and 12-month-olds habituated after an average of 76.48 s (*SD*=23.13 s) and also viewed about 7 habituation events (*M*=6.69, *SD*=2.30). There was no significant difference in the length of habituation or the number of videos observed, both *t*s<1.3, *p*>.19, thus confirming that both sets of participants were exposed to the same amount of input during habituation.

Potential order effects of the test events were examined with an Age (8-month-olds, 12-month-olds) by Order (first, second, third, fourth) ANOVA. As expected, there was a main effect of Order, $F(3,90)=4.73$, $p>.01$, $\eta^2=.14$, due to longer looking times to the first test event (*M*=9.76 s, *SD*=6.72 s) then the third (*M*=6.25 s, *SD*=6.11 s) and fourth test events (*M*=5.57 s, *SD*=3.19 s), both *p*s>.01. Differences between the first test event and the second test event (*M*=7.23 s, *SD*=6.47 s) were not significant, *p*=.08. Also, there were no differences in looking times for the second, third, and fourth test events. Thus, infants in both groups recognized the novel exiting event during the experiment and did so early in the testing phase.

To explore the influence of perceptual and probabilistic cues in the evidence mean looking times were submitted to a mixed ANOVA with Age (8 month-olds, 12-month-olds) as a between subjects variable and Probability (High, Low) and Ball location as within subjects variables (see Figures 2 and 3). This analysis revealed no significant effects or interactions, all *F*s<1.50, *p*s>.22. The mean looking times for both age groups are represented in Figures 2 and 3.

These results indicate that neither group generated expectations about single future outcomes when the samples included a ratio of 2:1. The decrease in looking times for the later test events suggests infants were not confused by the novelty of the test events. Moreover, these results are inconsistent with the prediction that a smaller and less disparate sample would support the use of probabilistic data to generate expectations for 8-month-olds. Finally, coupled with the results from Experiment 1, these results for 12-month-olds suggest that the expectation that a sample of evidence will yield the most likely future event appears to be mediated by the ratio of low-to-high probability items in the available evidence.

Experiment 3

The results from Experiments 1 and 2 indicate that 12-month-olds will generate expectations from samples of evidence with a 4:1 ratio, but not from samples with a 2:1 ratio. Moreover, 12-month-olds expected a high probability outcome when given a 4:1 ratio, whereas 8-month-olds expected a perceptually congruent outcome. However, it remains possible that 8-month-olds will generate expectations consistent with probabilistic data when evidence is more disparate – for example, when the ratio of one ball color to another is higher than 4:1 – because such samples may make more salient the most likely outcomes. Experiment 3 was designed to replicate the design of the previous experiments with the exception that the evidence in each box contained a ratio of 6:1.

Note that increasing the ratio difference requires increasing the sample size. There is support in the number discrimination literature for the conclusion that young infants (e.g., 6-month-olds) are better able to detect ratio differences for large sets of items (e.g., 8 vs. 4) than small sets of items (e.g., 4 vs. 2; Xu, 2003). Thus, this larger and more disparate sample may elicit probability-based expectations in younger infants. However, the presence of stronger cues to support discrimination between low and high probability items does not guarantee infants will generate expectations based on the probabilistic data. In fact, it is also possible this more disparate sample will intensify the perceptual contrast between the low and high probability balls and thus draw attention to the low probability item. This latter possibility suggests that this more disparate ratio of evidence could elicit the expectation that the samples should yield perceptually congruent events. Experiment 3 tested these predictions.

Method

Participants—The participants were 16 healthy full-term eight-month-olds (*M* =8 months, 17 days; range=7:20 to 9:10; 8 males, 8 females) and 16 healthy full-term twelve-month-olds (*M*=12 months, 9 days; range=12:3 to 13:1; 8 males, 8 females). Data from an additional 9 infants were excluded from the final sample due to fussiness (3), experimental error (4), and for failing to meet the criterion for inclusion in the data analysis (2). Methods of recruitment and reimbursement were the same as in the other experiments reported here.

Stimuli, design, apparatus, and procedure—The habituation and test trials involved events with seven balls, and the boxes were designed such that the ratio of high to low probability balls was set to 6:1. During habituation the duration of ball drops was set to 5 s. However, the time was divided equally such that each ball drop lasted approximately .714 s. In all other respects the experiment was identical to the previous experiments.

Results and Discussion

The first analysis considered the length of time to reach habituation. Eight-month-olds reached habituation in 102.73 s (SD=38.4 s) and viewed approximately 7 habituation events (M=6.88, SD=1.96) whereas 12-month-olds habituated on average in 92.88 s (SD=35.67 s) and also viewed approximately 7 habituation events (M=7.31, SD=2.47). Paired comparisons revealed these scores did not differ between age groups, both ts<1.10, ps>.17, thus confirming that both groups were exposed to the same amount of input.

As expected, the Age by Order ANOVA revealed a main effect of Order, $F(3,90)=7.26$, $p<.01$, $\eta^2=.19$, due to longer looking times to the first event (M=9.83 s, SD=6.67 s) than the second (M=6.31 s, SD=5.41 s), third (M=6.14 s, SD=4.91 s), and fourth test events (M=5.71 s, SD=5.01 s), all ps<.05. No other effects were significant.

Mean looking times for both age groups are represented in Figures 2 and 3. The main analysis revealed only an effect of Ball location, $F(1,30)=17.97$, $p<.001$, $\eta^2=.38$, due to longer looking times for Incongruent events (M=8.92 s, SD=3.56 s) than Congruent events (M=4.98, SD=1.88 s). As indicated in Figure 3 additional analyses confirmed this pattern of looking times among both groups: Twelve-month-olds, $F(1,15)=10.95$, $p<.01$, $\eta^2=.42$, and eight-month-olds, $F(1,15)=8.43$, $p < .05$, $\eta^2=.36$. No other main effects or interactions were significant, all Fs<1, ps>.38.

The results from Experiment 3 indicate that both groups of infants exhibited a perceptual-based looking pattern. Thus, unlike the 4:1 ratio, the 6:1 ratio caused 12-month-olds to expect a perceptually congruent outcome rather than a high probability outcome. Across the 4:1 and 6:1 conditions 8-month-olds consistently expected a perceptually congruent event.

General Discussion

The three experiments reported here provide several novel results concerning the ability of infants within the first year of life to use samples of probabilistic evidence to generate expectations about single future events. First, 8-month-olds did not expect the samples would yield the most probable outcomes, but they instead expected the samples would yield an event consistent with the location of items within the containers. Second, at least in one condition (e.g., 4:1), older infants expected the samples would yield the most probable outcome. Finally, 12-month-olds showed a probability-based looking pattern in the 4:1 condition, a perceptual-based pattern in the 6:1 condition, and showed no systematic pattern in the 2:1 condition, thus suggesting that the distribution of items within a sample affected which cues infants used to generate expectations about future events.

The results from these studies support the conclusion that the ability to generate expectations about future events undergoes development in the first year of life and that these changes may be due to the development of information-processing capacities. For example, probability-based expectations in the present task required the ability to represent the distributional properties of the evidence within the containers, while perceptual-based expectations required the ability to represent only the location of the ball closest to the exit of the containers. Because working memory capacity undergoes significant development in the first year of life (Perone, Simmering, & Spencer, 2011; Ross-Sheehy, Oakes, & Luck,

2003) it is possible younger infants were unable to encode all of the items in the evidence, but they were able to encode the location of the low probability balls because these items appeared earlier during habituation. Likewise, 12-month-olds may have showed the perceptual-based pattern in the 6:1 condition because they were unable to encode all of the items in this larger sample of evidence. Consistent attention to the location of the low probability ball might also suggest that 8-month-olds' attention was drawn to this item; attention to the low probability ball under these circumstances can be likened to the visual pop out effect (e.g., Adler & Orprecio, 2006; Quinn & Bhatt, 1998). From this perspective, the larger and more disparate sample could have intensified the perceptual contrast between the low and high probability items and thus caused 12-month-olds to focus on the low probability items. This interpretation may also explain why both groups of infants did not look longer at the perceptually incongruent events in the 2:1 condition: The perceptual contrast between the high and low probability items may not have been sufficient to draw attention to the low probability items.

These results suggest that 12-month-olds, but not 8-month-olds, are able to use multiple cues to generate an expectation about the most likely outcome from samples of probabilistic evidence. The ability to effectively use multiple cues in the evidence has been demonstrated as a later developmental achievement in several other domains (Sobel & Kirkham, 2007; Thiessen & Saffran, 2004; Wu, Mareschal, & Rakison, 2011). Teglas et al. (2011) also showed that 12-month-olds use multiple cues to generate expectations about samples of probabilistic evidence. In their study infants focused on ball location for shorter occlusion periods, but considered the statistical properties of evidence for the longer occlusion period, suggesting that occlusion time impacts infants' expectations about future outcomes. Results from the present work suggest that an additional feature of the available evidence, specifically the ratio of low-to-high probability balls in the sample, affects whether infants' generate expectations on the basis of the location of balls in the sample or the probabilistic information within the sample. Furthermore, these results indicate that the ability to consider both of these cues when they are presented simultaneously develops sometime after 8 months of age.

We do not dispute evidence from a large database of studies supporting the conclusion that infants are able to detect regularities from probabilistic data (Aslin, Saffran, & Newport, 1998; Kirkham, Slemmer, & Johnson, 2002; Saffran, Aslin, & Newport, 1996; Sobel & Kirkham, 2006). We also do not contest research that has shown infants younger than 8 months of age expect future events to maintain perceptual and physical cohesion (e.g., Kellman & Spelke, 1983). Instead, our view is that the ability to use samples of probabilistic data to generate expectations about future outcomes involves a set of skills, and that what develops between 8 to 12 months of age is the ability to effectively integrate these different skills. For instance, infants must have the information-processing capacities to represent the distributions in the evidence, be able to recognize that a sample of evidence can yield a single event, engage in some type of probability computation to determine the likelihood of future events, and assess the perceptual properties that afford a future event. The present results indicate that 12-month-olds use these skills when presented samples of probabilistic data and are able to modify their expectations based on specific features of the available evidence. Eight-month-olds may have the capacity to represent base rate information, but are unable to do so in the presence of conflicting perceptual cues or when the task involves processing input from more than one sample of evidence.

It is important to note several differences between the present tasks and those used in other research in this area. First, the present tasks included two samples of evidence while most other research presented only one sample (e.g., Xu & Garcia, 2008; Teglas et al., 2007; 2011). One notable exception is a recent study by Denison et al. (in press) who showed that when presented two containers of white and red balls with inverse distributions six-month-olds looked longer when a low probability sample was drawn from one of the containers than they did when a high probability sample was drawn from one of the containers. The present studies included a test event from each of two samples to assure infants did not attend exclusively to any single container. It could be argued that the additional processing demands of encoding input from the two samples exceeded the processing capacity of younger infants and thus made it difficult for them to represent the probabilistic evidence in the samples. Thus a simplified version of the task could be used to show that 8-month-old will rely on the distribution of balls to generate expectations about the samples. Though this may be the case, the idea that older infants, but not younger infants, respond on the basis of statistical properties under some conditions supports our main proposal that there are developmental changes in the ability to represent and use statistical properties of evidence within a sample to generate expectations.

The present study is among a few within this area of research to use video images of lottery-type events (e.g., Teglas et al., 2007; 2011). Other work has presented sampling scenarios in which a human actor presents the population of balls and then infants observe a sampling event. The active-sampling method has been used to test hypotheses about infants' assumptions about the procedures that were used to select evidence from a sample of evidence (Gweon, Tenenbaum, & Shultz, 2010; Xu & Denison, 2009). One could argue that the results from Experiment 1 indicate that 12-month-olds assumed the test event was randomly selected from the sample, while the 8-month-olds did not have the same assumption. We find this interpretation of our data unlikely for several reasons. First, studies that have explored infants' sampling assumptions have utilized human actors who selected evidence from samples. Even if infants have assumptions about sampling procedures, it is not known whether such assumptions hold for cases in which human actors are not responsible for sampling, such as in the present studies. Second, the sampling-assumptions perspective needs to be reconciled with the finding that 12-month-olds responded differently based on the specific features of the available data. In our view, the most parsimonious interpretation of the present results is that the contents of the sample, specifically the disparity between high-to-low base rate items and the location of items in samples, was sufficient to elicit expectations about future events.

Finally, this study was the first to utilize a visual habituation paradigm while prior studies used a familiarization method. One of the key differences between these paradigms is that habituation tends to involve longer periods of exposure to the test items (e.g., Teglas et al. 2007 presented a population of balls once for a total of 13 seconds, while mean looking to test events across the three current experiments was over 70 seconds). A benefit of the habituation paradigm is that greater exposure to the stimuli supports encoding of all of the features in the evidence and thus increases our confidence that looking behavior was governed by a violation of expectations about specific features of the evidence (e.g., Hunter & Ames, 1988). However, one drawback is that the habituation method may introduce demands that are absent in the probabilistic problems posed to young learners on a daily basis; some of our inferences are based on samples of evidence about which we have limited information. The familiarization method is a good model for situations in which inferences are based on limited exposure to information. For this reason, we believe researchers in this area should continue to employ both of these methods as a mechanism for understanding how infants develop the ability to generate expectations about samples of probabilistic evidence.

These results are also consistent with a large body of evidence showing that certain conditions cause children and adults to ignore base rate information during decision-making (Davidson, 1995; De Neys & Vanderputte, 2011; Jacobs & Potenza, 1991; Kahneman &

Tversky, 1972; 1973). By the age of 5 years people exhibit reasoning biases (e.g., representativeness heuristic) that cause them to ignore population statistics (e.g., about the base rate of bank tellers) when the available evidence about an individual from the population includes cues (e.g., a person is labeled an "activist") that suggest the individual is representative of a low base rate sub-population within the sample (e.g., feminist bank tellers). Findings such as these appear to be in conflict with other work showing that young children and adults are quite good statistical learners and are able to engage in rational inference (Anderson, 1991; Xu & Tenenbaum, 2007). This paradox is best reconciled by understanding the context sensitivity of human reasoning; what constitutes generalizable evidence depends on which features of the evidence are most salient (Gigerenzer, Hell, & Blank, 1988; Gigerenzer & Hoffage, 1995). The results from the present study might be taken as evidence that sensitivity to the way evidence is presented has roots in infant cognitive development.

In sum, the results of the three experiments reported here suggest that the ability to generate expectations about future events undergoes significant development from 8 to 12 months of age. The challenge for research in this area is to detail the conditions under which infants are able to detect and use samples of evidence to generate expectations about future events (see Teglas et al., 2011 for an excellent example of such work). Indeed, in our view the ultimate goal of research is this area is to provide an outline of the developmental milestones that mark the growth of our basic ability to use stochastic evidence to generate expectations about the future.

Acknowledgments

This research was supported by an NIH postdoctoral training grant (T32 MH019102). We wish to thank members of the Infant Cognition Lab at Carnegie Mellon University, especially Dussy Yermolayeva and Rosanna Breaux, for their assistance with various aspects of this research. We would also like to extend our sincere gratitude to the guardians and infants who participated in this research.

References

Adler SA, Orprecio J. The eyes have it. Visual pop-out in infants and adults. Developmental Science. 2006; 9:189–206. [PubMed: 16472320]

Anderson JR. The adaptive nature of human categorization. Psychological Review. 1991; 98:409–429.

Aslin RN, Saffran JR, Newport EL. Computation of conditional probability statistics by 8-month-old infants. Psychological Science. 1998; 9:321–324.

Bremner JG, Slater AM, Mason UC, Spring J, Johnson SP. Trajectory perception and object continuity: Effects of shape and color change on 4-month-olds' perception of trajectory identity. Developmental Psychology. in press.

Clearfield MW, Mix KS. Number versus contour length in infants' discrimination of small visual sets. Psychological Science. 1999; 10:408–411.

Davidson D. The representativeness heuristic and the conjunction fallacy effect in children's decision making. Merrill-Palmer Quarterly. 1995; 41(3):328–346.

Denison S, Xu F. Twelve- to 14-month-old infants can predict single-event probability with large set sizes. Developmental Science. 2010a; 13:798–803. doi: 10.1111/j.1467-7687.2009.00943.x. [PubMed: 20712746]

Denison S, Xu F. Integrating physical constraints in statistical inference by 11-month-old infants. Cognitive Science. 2010b; 34:885–908. [PubMed: 21564238]

Denison S, Reed C, Xu F. The emergence of probabilistic reasoning in very young infants: Evidence from 4.5- and 6-month-olds. Developmental Psychology. in press.

De Neys W, Vanderputte K. When less is not more: Stereotype knowledge and reasoning development. Developmental Psychology. 2011; 47:432–441. [PubMed: 21142361]

Gigerenzer G, Hell W, Blank H. Presentation and content: The use of base rates as a continuous variable. Journal of Experimental Psychology: Human Perception and Performance. 1988; 14(3): 513–525. doi:10.1037/0096-1523.14.3.513.

Gigerenzer G, Hoffrage U. How to improve Bayesian reasoning without instruction: Frequency formats. Psychological Review. 1995; 102(4):684–704. doi: 10.1037/0033-295X.102.4.684.

Gweon H, Tenenbaum JB, Schulz LE. Infants consider both the sample and the sampling process in inductive generalization. Proceedings of the National Academy of Sciences. 2010; 107:9066–9071.

Hunter, MA.; Ames, EW. A multifactor model of infant preferences for novel and familiar stimuli. In: Rovee-Collier, C.; Lipsitt, LP., editors. Advances in infancy research. Vol. 5. Ablex; Norwood, NJ: 1988. p. 69-95.

Jacobs J,E, Potenza M. The use of judgment heuristics to make social and object decisions: A developmental perspective. Child Development. 1991; 62:166–178.

Johnson SP. Development of visual perception. Wiley Interdisciplinary Reviews: Cognitive Science. 2010 doi: 10.1002/wcs.128.

Kahneman D, Tversky A. Subjective probability: A judgment of representativeness. Cognitive Psychology. 1972; 3:430–454.

Kahneman D, Tversky A. On the psychology of prediction. Psychological Review. 1973; 80:237–251.

Kirkham NZ, Slemmer JA, Johnson SP. Visual statistical learning in infancy: evidence of a domain general learning mechanism. Cognition. 2002; 83:B35–B42. [PubMed: 11869728]

Perone S, Simmering VR, Spencer JP. Stronger neural dynamics capture changes in infants' visual working memory capacity over development. Developmental Science. 2011; 14:1379–1392. [PubMed: 22010897]

Quinn PC, Bhatt RS. Visual pop-out in young infants: Convergent evidence and an extension. Infant Behavior & Development. 1998; 21:273–288.

Ross-Sheehy S, Oakes LM, Luck SJ. The development of visual short-term memory capacity in infants. Child Development. 2003; 74:1807–1822. [PubMed: 14669897]

Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. Science. 1996; 274:1926–1928. [PubMed: 8943209]

Sobel DM, Kirkham NZ. Bayes nets and babies: Infants' developing statistical reasoning abilities and their representation of causal knowledge. Developmental Science. 2007; 10:298–306. [PubMed: 17444971]

Kellman PJ, Spelke ES. Perception of partly occluded objects. Cognitive Psychology. 1983; 15:483–524. [PubMed: 6641127]

Starkey P, Cooper RG. Perception of numbers by human infants. Science. 1980; 210:1033–1035. [PubMed: 7434014]

Teglas E, Girotto V, Gonzalez M, Bonatti L. Intuitions of probabilities shape expectations about the future at 12 months and beyond. PNAS. 2007; 104(48):19156–19159. [PubMed: 18025482]

Teglas E, Vul E, Girotto V, Gonzalez M, Tenenbaum JB, Bonatti LL. Pure reasoning in 12-month-old infants as probabilistic inference. Science. 2011; 332:1054–1059. [PubMed: 21617069]

Tenenbaum JB, Griffiths TL. Generalization, similarity, and Bayesian inference. Behavioral and Brain Sciences. 2001; 24:629–641. [PubMed: 12048947]

Thiessen ED, Saffran JR. Spectral tilt as a cue to word segmentation in infancy and adulthood. Perception & Psychophysics. 2004; 66:779–791. [PubMed: 15495903]

Wu R, Mareschal D, Rakison DH. Attention to multiple cues during spontaneous object labeling. Infancy. 2011; 16:545–556.

Xu F. Numerosity discrimination in infants: Evidence for two systems of representations. Cognition. 2003; 89:B15–B25. [PubMed: 12893126]

Xu F, Denison S. Statistical inference and sensitivity to sampling in 11-month-olds. Cognition. 2009; 112:97–104. [PubMed: 19435629]

Xu F, Garcia V. Intuitive statistics by 8-month-olds. PNAS. 2008; 105(13):5012–5015. [PubMed: 18378901]

Xu F, Spelke ES. Large number discrimination in 6-month-old infants. Cognition. 2000; 74:B1–B11. [PubMed: 10594312]

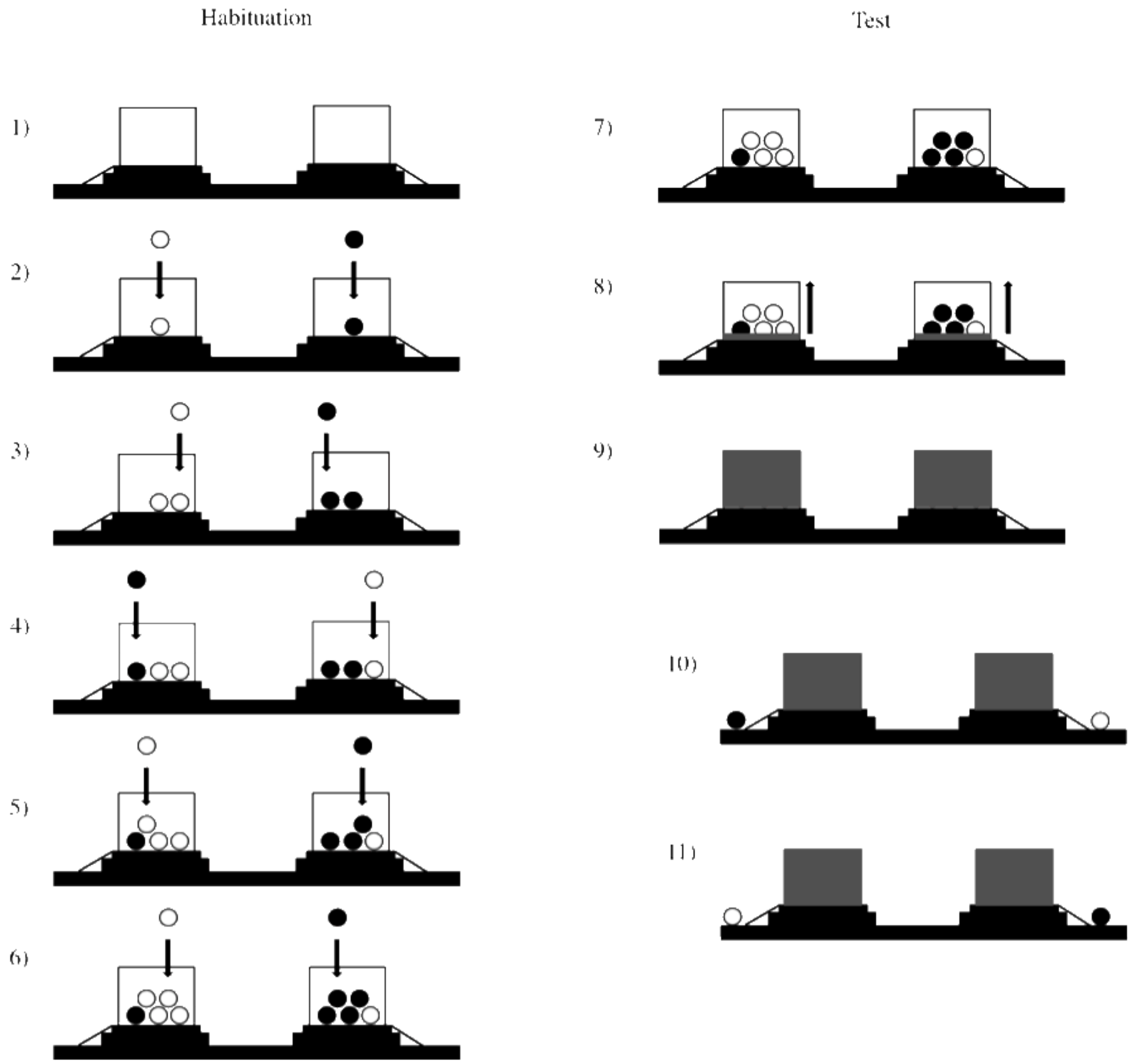

Figure 1.
Outline of the experimental design. During habituation infants were first shown the empty box (1) and then a series of events (2-6) during which a single ball was dropped into each of the boxes. During the Test phase the distributions of balls in each box were made visible (7) and then the contents of the box were shielded (8-9). Infants were then shown single test events that involved either a low probability event (10) or a high probability event (11) that were either the same color as the most adjacent ball in the container (i.e., perceptually congruent; as in panel 10) or a different color as the most adjacent ball in the container (i.e., perceptually incongruent; as in panel 11).

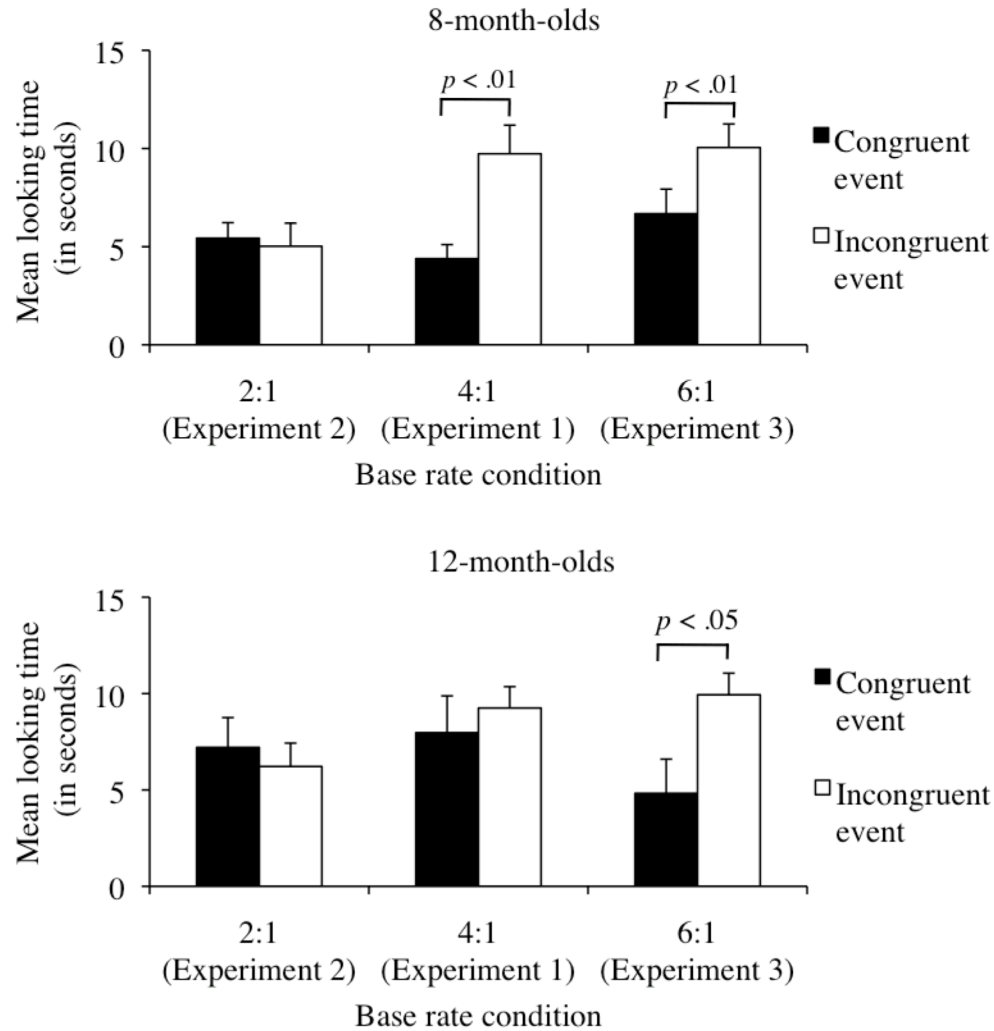

Figure 2.
Mean looking times (in seconds) to the Congruent event and the Incongruent event in Experiments 1-3 for 8- and 12-month-olds. Bars indicate one standard error from the mean.

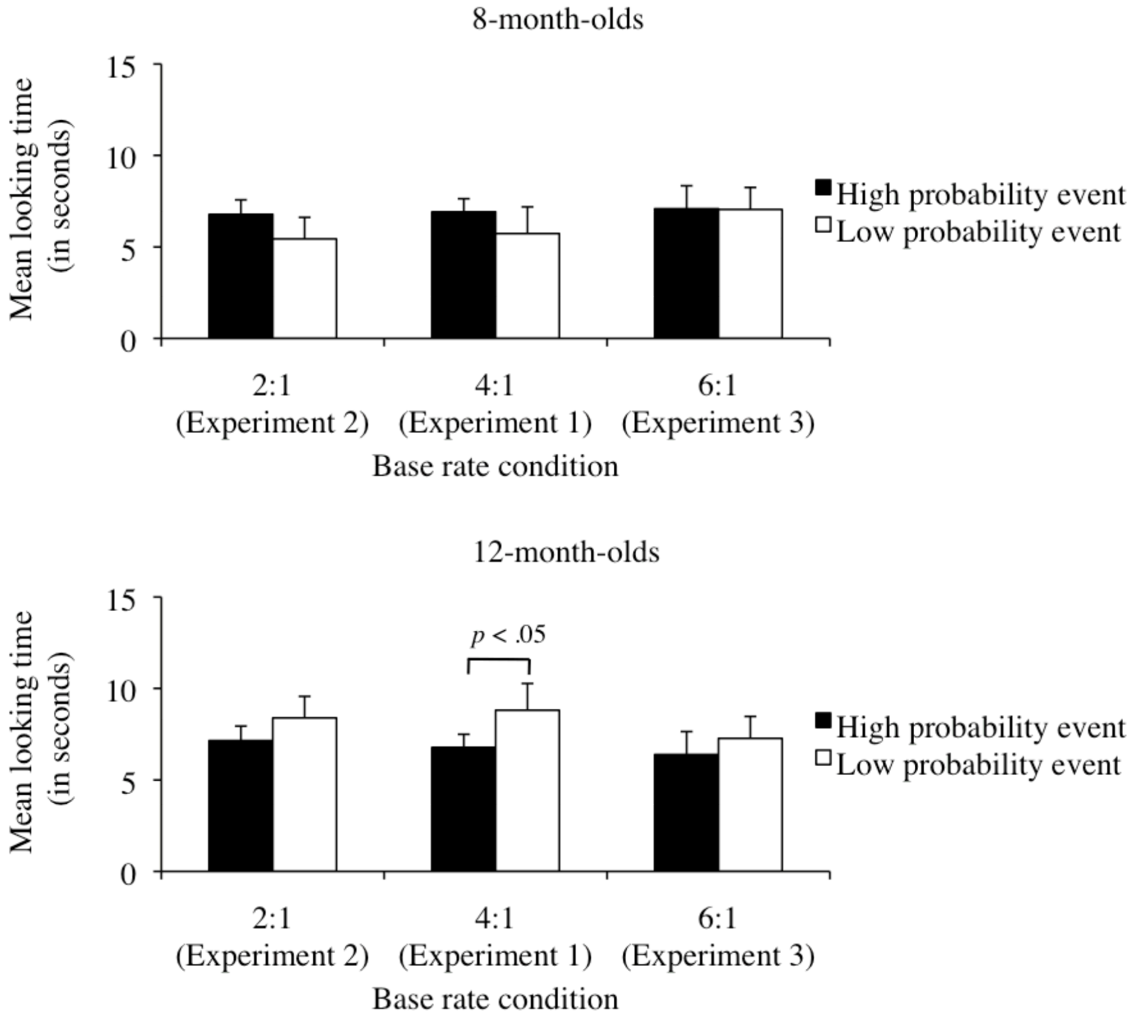

Figure 3.
Mean looking times (in seconds) to the High probability event and the Low probability event in Experiments 1-3 for 8- and 12-month-olds. Bars indicate one standard error from the mean.