Deficits in Top-Down Sensory Prediction in Infants At Risk due to Premature Birth

Highlights
- Prediction has been proposed to be essential for human development
- Infants at risk due to premature birth exhibit deficits in prediction
- Neural deficits were specific to prediction, and not to simple learning or perception
- Deficits were found early in development, suggesting a causal role for prediction

Authors
Lauren L. Emberson, Alex M. Boldin, Julie E. Riccio, Ronnie Guillet, Richard N. Aslin

Correspondence
lauren.emberson@princeton.edu

In Brief
Emberson et al. compare the neural capacity to predict upcoming stimuli in 6-month-old infants at low and high risk for developmental delays (full-term and preterm infants). Infants at high risk exhibited selective deficits in top-down sensory prediction, providing evidence that neural prediction supports human development.
Deficits in Top-Down Sensory Prediction in Infants At Risk due to Premature Birth

Lauren L. Emberson,1,2,4,* Alex M. Boldin,1 Julie E. Riccio,3 Ronnie Guillet,3 and Richard N. Aslin2

1Psychology Department, Princeton University, Peretsman-Scully Hall, Princeton, NJ 08544, USA
2Brain and Cognitive Sciences Department and the Rochester Center for Brain Imaging, University of Rochester, PO Box 270268, Meliora Hall, Rochester, NY 14627, USA
3Neonatology Unit, University of Rochester Medical Center, 601 Elmwood Ave, Rochester, NY 14642, USA
4Lead Contact
*Correspondence: lauren.emberson@princeton.edu
http://dx.doi.org/10.1016/j.cub.2016.12.028

SUMMARY

A prominent theoretical view is that the brain is inherently predictive [1, 2] and that prediction helps drive the engine of development [3, 4]. Although infants exhibit neural signatures of top-down sensory prediction [5, 6], in order to establish that prediction supports development, it must be established that deficits in early prediction abilities alter trajectories. We investigated prediction in infants born prematurely, a leading cause of neuro-cognitive impairment worldwide [7]. Prematurity, independent of medical complications, leads to developmental disturbances [8–12] and a broad range of developmental delays [13–17]. Is an alteration in early prediction abilities the common cause? Using functional near-infrared spectroscopy (fNIRS), we measured top-down sensory prediction in preterm infants (born <33 weeks gestation) before infants exhibited clinically identifiable developmental delays (6 months corrected age). Whereas preterm infants had typical neural responses to presented visual stimuli, they exhibited altered neural responses to predicted visual stimuli. Importantly, a separate behavioral control confirmed that preterm infants detect pattern violations at the same rate as full-terms, establishing selectivity of this response to top-down predictions (e.g., not in learning an audiovisual association). These findings suggest that top-down sensory prediction plays a crucial role in development and that deficits in this ability may be the reason why preterm infants experience altered developmental trajectories and are at risk for poor developmental outcomes. Moreover, this work presents an opportunity for establishing a neuro-biomarker for early identification of infants at risk and could guide early intervention regimens.

RESULTS AND DISCUSSION

The goal of this study was to establish a direct link between the neuro-cognitive impairments associated with prematurity and an infant’s ability to predict upcoming sensory input. To this end, we restricted our preterm population (born <33 weeks gestation) to those who did not experience severe medical complications or neurological insults and conducted the study at 6 months corrected age (i.e., matched to full-term infants based on time since conception, and not extra-uterine experience), before preterm infants missed any clinically identifiable developmental milestones. Testing at this young age allows us to circumvent the possibility that differences in prediction are arising from differences in developmental stage across the groups. Moreover, we employed a model task approach in which all infants receive equal experience with novel stimuli to control for possible differences in experience outside the lab.

Using functional near-infrared spectroscopy (fNIRS), a method for recording the hemodynamic response in the surface of the cortex using light [18–21], we recorded neural responses in 100 infants (50 preterm) to presented as well as predicted auditory and visual stimuli (see Figure 1). Following directly from findings in cognitive neuroscience (most closely [2, 22]), an important neural signature of top-down sensory prediction is responses to omitted information. If the developing brain is generating top-down predictions, an unexpected omission of visual information will result in activity in the same regions of the brain that process visual information. This has been observed in 6-month-old full-term infants: visually selective regions of the infant brain respond when visual input is unexpectedly omitted but exhibit no activity when the visual information was not expected to appear [5]. This paper extends this finding to infants at risk for poor developmental outcomes. To calculate the magnitude of the hemodynamic response, we averaged normalized changes in blood oxygenation from 5 to 9 s after stimulus onset within two neuroanatomically defined regions of interest (ROIs) (Figure 1; occipital: three NIRS channels; temporal: five NIRS channels; see Supplemental Experimental Procedures and [5, 23] for details on the MR-fNIRS coregistration method).

Prematurity Results in Differences for Predicted but Not Presented Sensory Input

Preterm and full-term infants exhibit the same pattern of response to presented auditory and visual stimuli. Building from [5], we examined the neural response of full-term infants during audiovisual trials in the temporal and occipital ROIs and confirmed the hypothesized perceptual cortex responses to auditory and visual stimuli: there was a significant increase

Please cite this article in press as: Emberson et al., Deficits in Top-Down Sensory Prediction in Infants At Risk due to Premature Birth, Current Biology (2016), http://dx.doi.org/10.1016/j.cub.2016.12.028
from baseline in both the temporal ($t(35) = 4.18$, $p < 0.001$, $d = 0.70$) and occipital ($t(35) = 4.74$, $p < 0.001$, $d = 0.79$) ROIs. Importantly, we found the same pattern of perceptual cortex responses in preterm infants: they exhibited a robust response in temporal and occipital ROIs during audiovisual trials (temporal: $t(42) = 4.57$, $p < 0.001$, $d = 0.70$; occipital: $t(42) = 3.13$, $p < 0.003162$, $d = 0.48$; Figure 2).

However, preterm and full-term infants differed in their responses to visual omission trials, in which the predicted visual stimulus is unexpectedly omitted. Full-term infants exhibited a robust occipital lobe response to the unexpected omission of a visual stimulus ($t(35) = 3.63$, $p < 0.001$, $d = 0.61$; Figure 2, right panel) that was statistically indistinguishable from the response of this region to the presentation of the visual stimulus in the audiovisual trials ($t(35) = 1.51$, $p = 0.14$, $d = 0.22$). Thus, we found patterns consistent with previous work [5] showing occipital lobe responses in 6-month-olds to the unexpected absence of a visual event (see [22] for a detailed investigation of these effects in adults using fMRI). However, in contrast, preterm infants exhibited a significant difference in occipital response levels between audiovisual trials and visual omission trials ($t(42) = 6.31$, $p < 0.001$, $d = 0.86$). This difference between preterm and full-term infants is also highlighted by the fact that there appears to be a negative occipital response during visual omission trials among preterm infants ($t(42) = -2.44$, $p = 0.01878$, $d = 0.37$). Thus, although preterm infants exhibit a robust response to an unexpected visual omission, the response is negative and significantly reduced from the responses of the same region to the presentation of a visual stimulus.

Directly comparing occipital cortex responses between preterm and full-term infants across audiovisual and visual omission trials (mixed ANOVA) reveals a significant interaction of trial type and birth status, $F(1,77) = 4.61$, $p = 0.035$, $\eta^2 = 0.04$, which is driven by a difference between preterm and full-term infants in occipital lobe response levels during visual omission trials, $t(55.60) = -4.36$, $p < 0.001$, $d = 1.03$, and a significant difference across trial types in preterm but not full-term infants. There were additionally main effects for birth status, $F(1,77) = 13.92$, $p < 0.001$, $\eta^2 = 0.15$, and trial type, $F(1,77) = 24.98$, $p < 0.001$, $\eta^2 = 0.23$. There was also a significant difference in occipital response during visual present trials between the two groups, $t(62.40) = -2.23$, $p = 0.0295$, $d = 0.52$.

Importantly, these findings are specific to the occipital cortex: as both trial types are initiated by the presentation of an identical, predictive auditory cue, no differences in temporal lobe activation are predicted. Indeed, there were no main effects for birth status, $F(1,77) = 2.40$, $p = 0.162$, $\eta^2 = 0.03$, or trial type, $F(1,77) = 1.02$, $p = 0.316$, $\eta^2 = 0.01$, and no significant interaction, $F(1,77) = 0.88$, $p = 0.351$, $\eta^2 = 0.01$. Both groups of infants exhibited a strong, positive temporal cortex response to visual omission trials (full-terms: $t(35) = 4.73$, $p < 0.001$, $d = 0.79$; preterms: $t(42) = 4.29$, $p < 0.001$, $d = 0.65$; Figure 2, left panel).

There are numerous explanations for negative hemodynamic responses in the fMRI literature. First, it may be that the negative response we observe when preterm infants experience an unexpected visual omission reflects a suppression of neural activity below spontaneous or baseline levels [24]. This interpretation of the negative BOLD response would result in the conclusion that preterm infants exhibit a neural pattern distinct from that of full-term infants when no expectation is present (control study [5]). However, an alternative explanation for negative BOLD responses is that there are differences in baseline response (e.g., [25]). In this case, a response to baseline stimuli in preterm infants could be elevated when compared to the full-term baseline. As all responses were recorded relative to baseline, an elevated baseline response would explain the smaller visually evoked response to the audiovisual trials and could also explain a significant reduction to the unexpected visual omission trials. In the Supplemental Information, we conducted a baseline correction that equates the level of neural response in the occipital ROI to the audiovisual trials across preterm and full-term infants and found the same effects across trial type and group.

**Figure 1. Task Structure and MR Coregistration of fNIRS Recordings**
(A) Overview of task structure. All trials began with a predictive sound/auditory stimulus. In the majority of the trials, this was followed by the predicted visual stimulus (audiovisual trials, right branch). However, in a minority of the trials (20% of trials after initial familiarization), this predictive sound was followed by an unexpected omission of the visual stimulus (visual omission trial, left branch).
(B) MR coregistration of fNIRS recordings was used to create two neuro-anatomically defined regions of interest (ROIs): temporal and occipital lobe ROIs (left and right images, respectively).

**There Are No Differences across Levels of Prematurity**
Preterm infants in this study were born from 23 to 32 weeks gestation. While all of these infants were born before the third trimester, the level of neural maturity at birth varied widely across this sample. Moreover, many risk factors for prematurity are much more severe in, or are restricted to, infants born extremely premature (<28 weeks gestation, [26, 27]). We investigated whether the deficits we observed are modulated by gestational age at birth: is there evidence for gradation in these deficits of top-down prediction across gestational age, or are these deficits uniform across infants born before the third trimester? First, removing early preterm infants from our sample (gestational age < 28 weeks) did not change the significance of any relevant statistical analysis. The remaining preterm infants showed significant occipital response during both audiovisual trials, $t(33) = -2.56$, $p = 0.01519$, $d = 0.02$, and visual omission trials, $t(33) = -2.57$, $p = 0.01487$, $d = 0.44$, with a significantly
negative response to the visual omission trials. There was still a significant difference in preterm occipital response between both trial types, \( t(33) = 5.46, p < 0.001, d = 0.87 \), and a significant difference in occipital response during visual omission trials between preterm and full-term infants, \( t(55.34) = -4.43, p < 0.001, d = 1.04 \). The fact that our results withstood the exclusion of this group indicates that the effects observed are not driven by infants born very early, but rather are effects that may distinguish preterm infants in general from full-term infants. Since relatively few \( (n = 9) \) infants fell into this early category, we do not have the statistical power to determine the specific influence of these extremely premature infants on our analysis. We also considered whether gestational age has a more subtle but graded effect on differences of responses to unexpected visual omissions. As Figure 3 illustrates, gestational age, within the preterm sample, does not account for variation in the neural response to visual omissions, \( r^2 = 0.01, F(1,41) = 0.47, p = 0.4949 \). Importantly, we also did not find that gestational age within the preterm sample explains occipital lobe responses to audiovisual trials when a visual stimulus is presented, \( r^2 = 0.01, F(1,41) = 0.56, p = 0.4581 \). Future work will address this surprising finding: is there is a categorical shift in early top-down prediction abilities after the first trimester, as the current data suggest, or would a sample including more infants born extremely premature reveal gradations in this ability?

**Socioeconomic Status and other Demographic Differences Do Not Explain the Effects of Prematurity**

In addition to prematurity, there were a number of demographic differences between our groups. Notably, our preterm sample had a lower socioeconomic status and were much more likely to have come from multiple births (e.g., twins, triplets). Importantly, additional analyses confirmed that the deficits in top-down sensory prediction observed across these groups were not explained by these other demographic differences. See Supplemental Experimental Procedures for more details.

**Preterm and Full-Term Infants Equally Detect Visual Omissions**

We found that infants born prematurely exhibit a significant reduction in the neural signature of top-down prediction. An alternate explanation of this result is simply that the unexpected visual omission is less unexpected to preterm infants. This could arise, for example, from reductions in cross-modal associative learning \([28]\). If preterm infants are slower to learn the cross-modal association between the sound and the visual event, that could explain the lack of occipital response to the visual omission trials. The current study was carefully designed to avoid this possibility (i.e., by providing infants with an equal amount of exposure to the stimuli in a paradigm that reduces learning demands through temporal overlap between audio and visual stimuli). Moreover, we previously examined neural responses to visual omissions for a control group of full-term infants who did not learn the audiovisial association and found that the occipital lobe did not respond differently from baseline, in contrast to the strong negative response observed in the occipital lobe for preterm infants \([5]\). Although predictive processes are well established as being involved in reinforcement learning, the nature of these signals is distinct from the type of prediction being studied here. Specifically, prediction errors involved in reinforcement learning are found in the basal ganglia and other subcortical circuitry \([29–31]\). This type of prediction is well reflective of a feedforward system in which this subcortical circuitry modulates expectations based on sensory input that continues on to modulate motor responses. By contrast, the current study investigates top-down prediction of sensory input that modulates perceptual cortices \([22, 32]\). Top-down sensory prediction is a distinct but complementary type of prediction that relies on the formation of an association (in the case of the current study) as an origin for a feedback signal that affects the perceptual cortices. Thus, while associative or reinforcement learning is likely involved in the current study, we aimed to isolate the effects of prematurity to top-down prediction signals, and not differences in associative learning, which can rely on a largely feedforward network architecture.

To confirm that preterm and full-term infants detect visual omissions similarly, we conducted a behavioral control experiment. A new sample of 50 full-term and 50 preterm infants were recruited using the same methods and populations as the fNIRS study. Specifically, we asked whether the exposure that infants received in the fNIRS study would result in similar looking preferences (i.e., length of looks) to visual omissions for preterm and full-term infants. Systematic looking preferences are canonically interpreted in relation to the strength of internal representations \([33]\). Thus, similar looking-time preferences would suggest that preterm and full-term infants detect the visual omission equally. After exposure to audiovisual pairs in an exposure similar to the fNIRS experiment, infants were presented with sequences of
familiar audiovisual trials in counterbalanced order with sequences that contained 50% visual omission trials. We found strikingly similar looking-time preferences across the two groups. Indeed, direct comparisons between the groups yielded no significant difference (see Figure 4 and Supplemental Experimental Procedures for full experimental details). This control experiment confirmed that there were no differences in detection of the visual omission trials across the two groups. This finding suggests that the differences between preterm and full-term infants are specific to top-down sensory prediction and do not arise from differences in foundations necessary to this task.

Why does being born prematurely disrupt the underlying neural mechanisms of top-down prediction? Having established that detecting visual omissions, medical complications, and socio-economic status do not explain these deficits, one possibility is that preterm infants’ early extrauterine experience negatively affects the development of this ability. Specifically, extrauterine experience is richly endowed with a myriad of patterns and statistical information that preterm infants receive well before full-term infants, during the third trimester, which is crucial for neural development (e.g., the development of long-range functional connectivity [34]). Receiving this experience too early may be detrimental because the developing brain is not ready for the input (e.g., neural connectivity is limited, which could prevent or bias learning). Another possibility, which is not mutually exclusive, is that the type of extrauterine experience that preterm infants receive is importantly different from that of full-term infants (e.g., due to necessary medical interventions). Numerous studies have investigated preterm infants with the goal of determining whether development of different abilities is supported by experience or neural maturation. In contrast to the present study and the large literature documenting the developmental difficulties associated with prematurity, many of these studies have concluded that preterm infants are relatively unimpaired or even accelerated in their development (e.g., [35–37]). It may be that relatively low-level early-developing abilities are not disrupted by premature birth, but that the disruption of foundational developmental mechanisms, as reported here, has developmental consequences that emerge later in life or in different domains.

In summary, we investigated neural signatures of top-down sensory prediction in young infants who are at risk for poor developmental outcomes due to premature birth. In comparison to their full-term peers, preterm infants exhibited typical neural responses to presented auditory and visual stimuli but showed substantially reduced neural responses to predicted visual stimuli. Moreover, these neural differences were present before infants missed any clinically apparent developmental milestones, suggesting that alterations in top-down sensory prediction could give rise to the developmental impairments that preterm infants experience but that are revealed months or years later (e.g., language delays, learning disabilities). This result dovetails with the finding that premature birth affects information processing and memory that predict cognitive outcomes later in life [38–40]. Overall, this work provides evidence that top-down prediction is part of the engine driving development and an important component of how the infant brain uses experience to mature. If being born prematurely affects the mechanisms by which development proceeds, as this work suggests, this would explain why the effects of prematurity are ongoing and compounding [14]. Moreover, this discovery presents an opportunity for establishing a neuro-biomarker to identify infants at risk and
to guide early intervention attempts once it has been established that these early life deficits predict poor developmental outcomes.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, one table, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2016.12.028.

AUTHOR CONTRIBUTIONS


ACKNOWLEDGMENTS

This study was approved by the Institutional Review Board of the University of Rochester. This work was supported by NICHD grants K99 HD076166-01A1 and R00 4R00HD076166-02, CIHR postdoctoral fellowship 201210MFE-290131-231192 to L.L.E., and NSF EAGER grant BCS-1514351 to R.N.A. Many thanks to Holly Palmeri and Kelsey Spear for their amazing help in the Rochester Baby Lab; to Joan Mertzbach, without whose recruitment prowess this project never would have happened; to Kelsey Jackson in the Princeton Rochester Baby Lab; and to all the infants and their caregivers who participated, especially the amazing NICU graduates!

Received: June 9, 2016
Revised: November 8, 2016
Accepted: December 12, 2016
Published: January 26, 2017

REFERENCES


Supplemental Information

Deficits in Top-Down Sensory Prediction in Infants At Risk due to Premature Birth

Lauren L. Emberson, Alex M. Boldin, Julie E. Riccio, Ronnie Guillet, and Richard N. Aslin
Figure S1: Distribution of occipital ROI responses to unexpected visual omissions for full-term and preterm infants. Distributional means presented in color coded dotted lines. Related to Figure 2.
Figure S2: Illustration of the effect of our baseline corrections on response during both trial types, audiovisual and visual omission. The solid black curves represent individual time-courses from each single trial type for full-term infants. The solid orange curves represent individual time-courses from each single trial type for preterm infants. Preterm mean responses were shifted to be the equivalent to the audiovisual trial for full-term infants (left panel: solid orange → dotted orange line). It is clear from this shift that it would arise from an elevated baseline response (i.e., the dotted orange line begins at a much higher level than full-terms). Making the corresponding shift in the visual omission trials (right panel: solid orange → dotted orange line) removes the “negative” response and instead reveals a highly attenuated response compared to the full-term infants. Related to Figure 2.
Figure S3: Mean levels of oxygenated hemoglobin during visual present and visual omission trials in (left to right) temporal lobe and occipital lobe. Preterm response was adjusted such that the mean response during visual present trials was the same for both preterm and full-term infants. See Figure S2 for illustration of the logic behind the baseline shift. Related to Figure 2.

Figure S4: Distribution of gestational age versus average oxygenated hemoglobin levels in the occipital ROI during audiovisual trials (visual omission trials shown in main text, Figure 3). Linear fit to preterm data is shown. All full-term infants had gestational age of 36 weeks or more. Related to Figure 3.
Figure S5: Comparison of responses to audiovisual and visual omission events in individual channels within the occipital ROI. Left panel shows occipital channels in full-term infants, right panel shows the same channels in preterm infants. Single channel results match overall occipital ROI results indicating no spatial selectivity of these effects. Related to Figure 2.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Estimator</th>
<th>$t(35)$</th>
<th>$p$</th>
<th>$F(1, 35)$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication</td>
<td>-1.07</td>
<td>-2.52</td>
<td>0.0165</td>
<td>6.34</td>
<td>0.15</td>
</tr>
<tr>
<td>Daily Living</td>
<td>-1.21</td>
<td>-3.32</td>
<td>0.00214</td>
<td>10.99</td>
<td>0.24</td>
</tr>
<tr>
<td>Socialization</td>
<td>-0.54</td>
<td>-2.07</td>
<td>0.0459</td>
<td>4.28</td>
<td>0.11</td>
</tr>
<tr>
<td>Motor Skills</td>
<td>-0.79</td>
<td>-4.84</td>
<td>&lt;0.001</td>
<td>23.41</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Table S1: Results of linear regressions modeling the relationship between SES and each of the four Vineland-II sub-domains. All show SES significantly predicting scores in each domain. Related to Figure 3.
Supplemental Experimental Procedures

Participants

Preterm infants, born at 32 weeks gestation or less, were recruited through the Neonatal Intensive Care Unit (NICU) at the University of Rochester Medical Center. Strict exclusionary criteria were applied to this population to attempt to isolate the effects of prematurity on neural and cognitive development and not other known medically-based risk factors. In other words, infants were excluded if they had risk factors for poor cognitive development beyond being born prematurely. Specifically, preterm infants were excluded from the study if they met any of the following exclusionary criteria: intraventricular hemorrhage (IVH, grade 3 or 4), periventricular leukomalacia (PVL), severe bronchopulmonary dysplasia (BPD, defined as infants who require supplementary oxygen after discharge), major surgeries, seizures, failing their hearing screening, known chromosomal abnormalities, major malformations, congenital viral infections, retinopathy of prematurity (ROP) requiring intervention (i.e., laser surgery) and weight and head circumference less than the 10th or greater than the 90th percentiles at birth. Researchers approached families in the NICU if their infants met the criteria for inclusion to the study, had not indicated that they didn’t wish to be approached for research studies and with the permission of their attending physician. Researchers explained the study and asked for consent to contact the family after discharge and closer to the time of the study. These interested families were then contacted and, if responsive and interested, scheduled for a study when their infant(s) were 5-7 months corrected gestational age. Full-term infants were recruited through the database of interested families for the Rochester Baby Lab. Inclusion criteria for full-term infants were birth at 36 weeks gestation or later, no major health problems or surgeries, and normal vision and hearing. Full-term infants participated in the study between 5-7 months of age.

Overall, there were 100 infants recruited for the study (50 full-term and 50 preterm), 20 were excluded (6 preterm, 14 full-term) due to poor signal quality and failing to watch the video for the required amount of time. One preterm infant was found to have not met inclusion criterion after recruitment (head size <10th percentile at birth). Of the 79 who comprise the final sample for the study, 36 were full-term and 43 were preterm. Note that 26 of the full-term infants comprise the visual omission experiment reported in Emberson et al (2015)[S1]. There were no racial, $c^2(8, N = 79) = 11.73, p = .1637$, or sex, $c^2(1, N = 79) = .11, p = .7422$, differences between the preterm and full-term subjects. In total, there were 62 white, 5 black, 10 other (e.g., multiple races), and 2 infants whose race was not reported. Four infants were reported as Hispanic and 75 were non-Hispanic. There were 39 female and 40 male infants.

Gestational age of the preterm infants ranged from 23 weeks to 32 weeks. However, the preterm sample skews towards the latter end of the gestational range reported: mean = 30.01, median = 30.86 gestational age in weeks. In fact, only 9 infants (21%) were born at less than 28 weeks gestation. Thus, the majority of infants in the preterm sample were born between 28 and 32 weeks age.
gestation making them very but not extremely premature. All of the preterm infants passed their hearing screening. No infants had laser surgery for ROP and, at their last known visit to the optometrist, all but 4 had fully mature retinas. Of those 4 (9.30%), all had zone II or III, stage I or below ROP. Five included infants (11.53%) had IVH, three with grade 1 and two with grade 2. All five had unilateral IVH, with one infant having the right side of the brain affected. Eleven infants (25.58%) were discharged on medications for bronchopulmonary dysplasia (BPD). Seven infants (16.28%) were rehospitalized for BPD, with four of these infants having not been discharged with medications for BPD. Four included preterm infants (9.30%) were exposed to prenatal substance abuse (including tobacco and alcohol use) but had no known health issues resulting from this exposure. It is not known how many full-term infants were exposed to prenatal substance abuse.

0.1. Design

We sought to test whether there are differences in the neural signatures of visual prediction between typically developing full-term infants and preterm infants who are at-risk for poor developmental outcomes. To this end, the full-term and preterm infants recruited for the study had identical experimental procedures: both groups were given experience where a sound predicts a visual event. After this exposure, infants were presented with trials where auditory events were followed by an unexpected visual omission. During all stimulus presentation, we employed fNIRS to record hemodynamic responses in the occipital and temporal cortex to provide a measure of the neural activity in these regions. The experiment continued until infants stopped watching consistently or became fussy. After the experiment was over, all families filled out questionnaires about their infants (e.g., race and ethnicity) and their family demographics (e.g., educational level of each parent and household income). The sole difference in experimental design across the groups is that caregivers of preterm infants filled out an additional standardized questionnaire designed to assess the developmental progress of their infant (Vineland II). Parents of full-term infants did not fill out the Vineland questionnaire.

0.2. Materials and Apparatus

0.2.1. Experimental Stimuli

All infants were presented with the same auditory and visual stimuli. These stimuli were played while the monitor presented a monochromatic grey screen with a white box (black bordered) in the middle. Auditory stimuli were novel, non-speech auditory stimuli, well described as a honk like from a clown horn and an unusual rattle sound. The visual stimulus, a red cartoon smiley face, was presented in two different ways: Entering the white box from either the top or the bottom, moving into the box to touch the opposite side of the box in 500ms and then exiting the box by traveling toward the same edge where it entered in 500ms. Both the auditory stimuli and the duration of the presence of the red smiley face were 1 second. All trials consisted of a combination of these
simple auditory and visual stimuli. During inter stimulus intervals (ISI), baseline stimuli were presented: dimmed fireworks video \textsuperscript{S2} and calming instrumental version of “Camptown Races” by Baby Music (album released 2010). Stimuli were presented on a Tobii 1750 eye tracker screen measuring 33.7 by 27 cm and computer speakers placed directly below the screen but behind a black curtain. Sounds were presented between 64 and 67dB. Stimuli were presented using MATLAB for Mac (R2007b) and Psychtoolbox (3.0.8 Beta, SVN revision 1245).

0.2.2. fNIRS Recordings

FNIRS recordings were gathered using a Hitachi ETG-4000 with twenty-four NIR channels: 12 over the back of the head to record bilaterally from the occipital lobe, and 12 over the left side of the head to record from the left temporal lobe. The channels were organized in two 3x3 arrays, and the cap was placed so that, for the lateral array, the central optode on the most ventral row was centered over the left ear and, for the rear array, the central optode on the most ventral row was centered between the ears and over the inion. This cap position was chosen based on which NIR channels were most likely to record from temporal and occipital cortex in infants \textsuperscript{S3}. Due to curvature of the infant head, a number of channels did not provide consistently good optical contact across infants (the most dorsal channels for each pad). We did not consider the recordings from these channels in subsequent analyses and only considered a subset of the channels (7 for the lateral pad over the ear and 5 for the pad at the rear array). FNIRS recordings were collected at 10 Hz (every 100 ms). Using a serial port, marks were presented from MATLAB on the stimulus presentation computer to the Hitachi ETG-4000. Marks were sent for the start and end of each presentation type for the given experiment (e.g., blocks of AV trials, single AV trials, single omission trials). Caregivers were not masked during fNIRS recordings because we have no reason to believe that any caregiver response to our experimental manipulations would result in selective neural responses in the regions of interest and, unlike behavioral or looking time studies where this practice is standard, we are not looking at subtle changes in behavior which can logically be affected by caregiver response. Moreover, an unmasked caregiver can aid the researcher by preventing the infant from grabbing at the NIRS cap and disrupting fNIRS recordings.

0.2.3. Questionnaires

All caregivers were asked to filled out questionnaires to determine the basic demographic information of their infant (e.g., race, ethnicity, confirming date of birth, whether they have siblings) and their family (e.g., education level of each parent, household income). Questionnaires were typically administered during the family’s visit to the lab and usually after the infant had participated in the fNIRS experiment. However, some families were asked the questions over the phone after their appointment if they were unable to fill out the questionnaire during their visit. While all families were requested to fill out the SES questionnaire, some declined to fill out the questionnaire or were not available by phone.
Others declined to fill out subsets of the questionnaire (e.g., family salary). In addition, caregivers of premature infants were asked to fill out the Vineland II: Vineland Adaptive Behavior Scales, Second Edition: Parent/Caregiver Rating Form. As this form is meant to assess developmental stage for children from 0 months to 18 years and the infants were 5-7 months CGA, caregivers were given a subset of this form to expedite the process of filling out the questionnaire. Caregivers were provided with written instructions and verbally instructed to fill out the form to the best of their ability and that it was not unusual for most of the form to be left blank as the questionnaire was meant to assess developmental stages well-beyond where their infant was. Caregivers were free to ask questions as a research assistant was always present during this time.

0.3. Procedure
0.3.1. Stimulus Presentation

Stimulus presentation always started with the presentation of a white box against a grey screen, and, immediately after, one of two auditory stimuli was presented. In the majority of trials, this auditory stimulus was followed by the appearance of the visual stimulus appearing at either the top or the bottom of the box. The visual stimulus was presented 750ms after the onset of the auditory stimulus and the auditory and visual stimulus overlapped for 250ms. Each sound was uniquely paired with one direction of movement for the smiley face, counterbalanced across infants. Individual pairs were employed with the same frequency throughout the experiment and in randomized order. This sub-pairing of audiovisual stimuli was intended to boost infant engagement and interest in the task but the main experimental question concerns the more general prediction of visual events following auditory cues. After the conclusion of the visual stimulus, the empty white box and the grey screen (static visual presentation with silence) was presented for 1000-1500ms (randomly determined). In the minority of trials, the visual stimulus was not presented. In these unexpected visual omission trials, the duration of the trial did not change but the period of time where the visual stimulus would have been presented consists of a static presentation of the white square. In other words, in trials where there is no visual stimulus presented, infants view the static white square for the duration of the trial. Thus, the majority of the time, infants saw audiovisual trials where an auditory stimulus is followed by the visual stimulus. However, in some trials, the auditory stimulus is not followed by the visual stimulus which is an unexpected visual omission trial.

These two trial types were either presented on their own as single-trials or organized into blocks. Blocks always consisted of 6 audiovisual trials: Three of each specific audiovisual pairing, randomly ordered and each separated by the jittered ISI (1-1.5 seconds). Based on the assumption that, with these novel auditory and visual stimuli, infants do not start the experiment with an expectation that these auditory stimuli will be followed by visual stimuli, the experiment began with the presentation of three blocks (18 audiovisual trials) to provide infants with some initial learning of the specific AV pairs and generate appropriate sensory expectations.
After the initial 3 block presentation, single trials were introduced and consisted of the presentation of a single audiovisual trial or a single visual omission trial. All neural data reported in this paper consist of direct comparison of single audiovisual trials and single visual omission trials. Specifically, after the 3 block familiarization, infants were presented with a random order of 1 block (6 audiovisual trials), 2 single audiovisual trials, and 2 single unexpected visual omission trials. This resulted in 1 block and 4 single trials that were presented repeatedly in a fully randomized order until the infant became fussy or consistently inattentive. This design ensured that 1) only 20% of the trials presented unexpected visual omissions in order to maintain expectations for a visual stimulus following an auditory stimulus and yet 2) provide equal numbers of trials and trial types to compare across audiovisual and visual omission trials (both single trials, presented with the same frequency and at consistent times throughout the experiment). In-between single trials and blocks, baseline stimuli were presented (dimmed fireworks video [S2] and calming instrumental version of “Camptown Races” Baby Music, album released 2010) for a jittered ISI of 4-9 seconds (mean = 6.5 seconds [S3]).

The experiment was conducted in a darkened room with dark floor-to-ceiling curtains surrounding the infant and their caregiver. Only the monitor (Tobii eye tracker) was visible to the infant as all other equipment (e.g., speakers, computers) was on the other side of the curtains and out of sight. Infant’s sat of their caregiver’s lap. Caregivers were instructed to not interfere with the infant’s watching of the video but to make sure that they did not grab at the cap on their head or rub up against them with the cap which would cause it to move. We also asked that they encourage the infant to be as still as possible but that they could allow the infant to move and stand up if it was necessary to keep the infant contentedly watching the video. The researchers watched the caregiver and infant from a video camera positioned beneath the monitor.

0.3.2. FNIRS Data Analysis

The raw data were exported from the Hitachi ETG-4000 to MATLAB (version R2015a for Mac) for subsequent analyses with HomER 2 (Hemodynamic Evoked Response NIRS data analysis GUI, version 1.5) using the default preprocessing pipeline of the NIRS data. First, the raw (intensity) data were converted to optical density. Next, a PCA filter and additional analyses are performed to identify and remove motion artifacts. The data was then low-pass filtered with a cutoff frequency of 3 Hz to remove noise and the modified Beer-Lambert law was used to determine the hemoglobin concentration for each channel (DOT.data.dc output variable was used for all subsequent analyses). A more detailed description can be found in the HomER 2 Users Guide [S5]. Timing information (mark identity and time received by the ETG-4000 relative to the fNIRS recordings) was also extracted from the ETG-4000 data using custom scripts run in MATLAB R2015a.

Subsequent analyses were conducted in MATLAB (R2015a) with custom analysis scripts. First, the continuous data was segmented and sorted into individual trial types based on the timing of marks. Because the experiment
was ended when the infant became inattentive or fussy, we excluded trials at the end of the experiment that were not presented past the mean duration of the baseline (duration of stimulus presentation + 6.5 seconds). The number of complete trials was determined for each trial type and it was evaluated whether the infant meet the inclusion criteria of watching a minimum of 4 single trials of both types (e.g., 4 audiovisual trials and 4 visual omission trials, see Participants for the number of infants excluded for not watching a sufficient amount of time for each study type). Full-term infants included in analysis looked on average for 6.86 single audiovisual trials (SD = 2.38, range = 2-12), 6.92 visual omission trials (SD = 2.32, range = 4-12), and 6.28 blocks (SD = 1.03). Included preterm infants looked on average for 7.65 single audiovisual trials (SD = 2.22, range = 3-12), 7.42 visual omission trials (SD = 2.06, range = 4-12), and 6.63 blocks (SD = 0.93).

Then, for each infant, the average concentration of oxygenated and deoxygenated hemoglobin per channel was determined for each condition. Infants were excluded at this point if the data collected was still noisy: a combination of visual inspection, experimental notes on optical contact and the presence of hair, and output from the otparex.m script which provided a measure of the number of “bad” channels was employed to determine signal quality for each infant. Importantly, the decision to include or exclude infants was made before group averages were determined and was not revisited in order to minimize experimenter bias. Then, the average and variance of responses for oxygenated and deoxygenated hemoglobin were determined within each ROI for each infant. A single analysis time window, 5-9 seconds after stimulus onset and defined a priori, was used for all trial types: five seconds was selected to be the point where the hemodynamic response typically begins to exhibit an increase in infants of this age based on numerous previous studies [S6] and to be conservative with regards to any residual response to the previous trial (5 seconds after stimulus onset is an average of 14 seconds after the stimulus onset of the previous trial). As is standard, we analyzed the response to the mean of the jittered ISI; our jitter interval was selected based on a validation study with visual stimuli using fNIRS [S4].

Correcting for Baseline Visual Responses Across Preterm and Full-term Infants

There are at least two possible explanations for the negative occipital response of preterm infants during visual omission trials. First, it may be true that preterm infants display an occipital response during visual omission that is negative or opposite to full-term response. However, it is also possible that response to baseline stimuli in preterm infants is elevated when compared to the full-term baseline (taken here to be zero response). An elevated baseline response would explain a reduced visually-evoked response to the audiovisual trials but could also explain a significant reduction to the unexpected visual omission trials.

Pursuing the possibility that preterm infants may have a different baseline response than the full-term infants, we adjusted the preterm responses such that the mean response during audiovisual trials was the same for preterm and
full-term infants in each of the regions of interest (ROIs, see Figs. S3 and S2). We then adjusted the response to the visual omission trials by this same amount. This method resulted in an increase in the occipital ROI by .00033 mm Mm which corresponds to a 128.89% increase in the mean audiovisual response and a 205.98% increase in the mean visual omission response. The temporal response was increased by .00026 mm Mm, corresponding to a 73.51% increase in mean audiovisual response and a 75.19% increase in mean visual omission response. This change, by design, eliminates the significant difference in response in both lobes during visual present trials between preterm and full-term infants. This is reflected in the absence of a main effect of birth status for occipital responses, $F(1, 77) = 1.10$, $p = .298$, $\eta^2 = .01$.

After this correction, preterm occipital response during unexpected visual omission becomes significantly positive, $t(42) = 2.59$, $p = .01311$ $d = .40$, but there is still a significant difference in response between visual present and visual absent trials in preterms. There is a marginally significant difference in occipital response during visual omission trials between full-term and preterm infants $t(55,60) = -1.91$, $p = .06102$, $d = .45$ and a significant interaction between trial type and birth status, $F(1, 77) = 4.61$, $p = .035$, $\eta^2 = .04$. Thus, taking into account possible differences in baseline responses across preterm and full-term infants, there is still evidence of differences specific to the visual omission trials, especially considering that the mean preterm signal was more than doubled in this correction. There is also still a main effect of trial type $F(1, 77) = 24.98$, $p < .001$, $\eta^2 = .23$.

0.4. There is No Evidence of Spatial Selectivity in Top-Down Prediction Deficits

We examined the possibility that there are spatial differences in visual activation and/or top-down prediction of visual input between preterm and full-term infants. To this end, we performed additional analyses on each of the three (3) individual channels placed within the occipital ROI (channels 3, 4, and 5). First, we investigated spatial selectivity in full-term infants. We found no significant differences in response between audiovisual and visual omission trials (the main piece of evidence for top-down, sensory prediction) in any channel (channel 3: $t(35) = .042$, $p = .676557$; channel 4: $t(35) = 1.88$, $p = .0686$; channel 5: $t(35) = 1.64$, $p = .10967$). This is consistent with our analysis of the overall mean response in the occipital ROI and demonstrates no spatial selectivity in these effects in full-term infants. The overall results for preterm infants also indicate that the same pattern observed in the overall ROI is observed in each channel. Specifically, each channel showed a significant difference in response between the two trial types (channel 3: $t(42) = 5.66$, $p < .001$; channel 4: $t(42) = 3.34$, $p = 0.001754$; channel 5: $t(42) = 6.05$, $p < 0.001$). It is of note that in each set of infants channel 4 seems to show a slightly weaker response than the other two channels: This may be because this channel is localized over the midline of the brain and tends to be noisier. See Figure S5.

It should be noted that these analyses are not corrected for individual variation in channel placement or head size. The use of an occipital ROI and averaging these three channels in the main analyses is an effort to increase the spatial
certainty and to examine large-scale neural patterns. However, the investigation of single channels provides a preliminary estimate of the spatial distribution of these broader scale signals. Given that these analyses are within subjects, it can be assured that the placement of the channel is the same across the comparisons, while spatial localization is less certain, but the between-subject analyses in the subsequent paragraph involve greater spatial uncertainty and specifically whether the channel location is the same across the full-term and preterm infants.

To provide convergent evidence, we compare responses in each channel between full-term and preterm infants. Overall, these results match the effects found in the broader occipital ROI as well. Channels 3 and 5 show significant differences between full-term and preterm audiovisual (channel 3: $t(77) = 2.08, p = 0.041184$; channel 5: $t(77) = 2.13, p = 0.036327$) and visual omission (channel 3: $t(77) = 3.73, p < 0.001$; channel 5: $t(77) = 3.64, p < 0.001$) responses. Responses in channel 4 do not show significant differences in either trial type (audiovisual: $t(77) = 1.47, p = 0.144937$; visual omission: $t(77) = 1.96, p = 0.053655$), however, as noted above, this channel has a consistently weaker response than the other two.

These results confirm our conclusion that there is no spatial selectivity of the preterm visual response during either trial type and differences in brain organization are likely not the explanation of our findings.

**Socioeconomic Differences Do Not Explain The Effects of Prematurity**

Socioeconomic status (SES) is a risk factor for prematurity [S7]. To quantify SES in our sample, we created a measure that represents a combination of average parental education and income: both important aspects of SES. As expected, full-term infants scored significantly higher than preterm infants in this measure, as indicated by a Wilcoxon Rank-Sum Test, $Z = -3.04, p = .002339$. To determine whether deficits in top-down, sensory prediction may be related to SES rather than directly to prematurity, we examined the relationship between SES and occipital response during both audiovisual and visual omission trials. Considering both groups together, we find that SES is a significant predictor of occipital lobe response in both audiovisual and visual omission trials, $bs > 3.28 \cdot 10^{-5}, ts(66) > 2.27, ps < 0.027$. Dividing both groups according to birth status, we find this relationship holds only for those born premature and again in both trial types $bs > 3.11 \cdot 10^{-5}, ts(37) = 2.14, ps < 0.0263$ (in full-term infants, $ps > 0.16$). It is important to note two things: 1) this relationship is not extremely robust and was not significant before the exclusion of 1 preterm from our sample (based on inclusionary criteria, head was small for gestational age at birth, but was erroneously included in the study because of the inclusion of their twin); 2) the full-term sample only includes relatively high SES families, while the preterm group contained both very high and very low SES families, making it difficult to determine whether the effects of SES are restricted to the preterm group or whether we are unable to see any differences in the full-term group because of our sampling bias. However, given the positive relationship between visual omission responses and SES in the preterm group, we sought to
establish that differences in top-down prediction between preterm and full-term infants remain after accounting for any effects of SES. To this end, we restricted our preterm sample to only families in the same SES range as the full-term infants (10 infants excluded because their family’s SES was below the minimum SES for the full-term sample, 13.5/23, see Experimental Methods). Even when both preterm and full-term infants have the same SES, there is a significant reduction in occipital lobe response to unexpected visual omission ($t(44.42) = -3.00, p=0.004$; wilcoxon signed rank test, $W=243, p=0.005$). Thus, while we find evidence that SES is a significant predictor of occipital lobe responses in preterm infants, SES does not account for the overall effect of prematurity because a difference is found for high SES infants only.

Accounting for Differences in Multiple Births across Preterm and Full-term Populations

More than half (51.16%) of our preterm sample were born as members of a multiple birth (twins, triplets and one group of quadruplets) as opposed to 2.86% of our full-term sample (one member of a twin). This striking difference in the populations introduces the possibility that differences between infants born as part of a multiple birth versus those born in a single birth is accounting for our results as opposed to prematurity per se. In order to examine whether the inclusion of the infants born as a part of a multiple birth may have altered our results, we examined the occipital lobe responses of only the infants born in single births. In the occipital lobe, most statistics maintained their significance. Full-term infants showed a robust response during both AV+, $t(34) = 4.51, p < .001$, and AV-, $t(34) = 3.41, p = .001674$, trials. There was no significant difference between visual present and visual omission response among the full-term infants, $t(34) = 1.41, p = .1664$. Preterm infants showed a robust response during AV+ trials, $t(20) = 2.39, p = .02662$, and there was a significant difference between trial types among preterm infants, $t(20) = 4.50, p < .001$. Additionally, there was a significant difference in response during AV- trials between preterm and full-term infants, $t(53.98) = -3.58, p < .001$. Two tests, however, did not retain their significance after infants born in multiple births were removed. There was no longer a difference between full-term and preterm infants in AV+ response, $t(48.06) = -1.30, p = 2.006$, and preterm infants no longer showed a robust negative response during AV- trials, $t(20) = -1.42, p = .171$. While this may indicate that there is an effect of multiple births on our results (e.g. infants born from multiple births explain the negative response), it is also possible that dramatically decreasing the sample size reduced our statistical power and thus caused the loss of significance. To account for this possibility, we performed a random resampling that excluded the same number of preterm and full-term infants (22 and 1, respectively). When this random resampling was performed 100 times it yielded average p-values that were also insignificant for the above tests. Specifically, there was no significant difference between full-term and preterm infants in AV+ response, $t(51.40) = -1.92, \bar{p} = .1055362$, and there was no significant preterm response during AV- trials, $t(20) = -1.66, \bar{p} = .1959855$. This indicates that the loss of significance is likely due to an increase in error
because of the reduced number of subjects rather than a real effect of multiple births.

We also explored whether the genetic link between infants born during the same multiple birth could have an effect on the results. Specifically, we addressed the concern that including multiple infants from a single pregnancy could be artificially augmenting our results by effectively repeating subjects. To address this possibility, we repeated our analysis with all infants but with all members of the same multiple birth averaged to create one combined subject. Among the 22 preterm infants born as part of a multiple there were 12 distinct births. This means that including only a single sample per birth would reduce our sample size by only 10 subjects. Of these 12 births, the data for seven of these births did result in more than one baby in our final sample, and composite values were created as an average of the infants born during each birth. Only one infant of each of the 5 remaining multiple births was included in our final sample, so they were included in this analysis without alteration. Since none of the full-term data was changed in this analysis, all statistics are identical to our original analysis. Changing our data to only include a single sample from each birth did not change any aspect of our analyses. As before, in the occipital ROI, full-term infants had robust response during both AV+ and AV- trials and there was no significant difference between the two trial types. Preterm infants, whose data was changed in this averaging, also showed robust response during both trial types (AV+: $t(32) = 2.85$, $p = .007655$; AV-: $t(32) = -2.19$, $p = .03607$), and there was a significant difference between them, $t(32) = 6.06$, $p < .001$. There was also a significant difference between preterm and full-term response during both AV+, $t(64.04) = -2.03$, $p = .04621$, and AV-, $t(58.10) = -4.24$, $p < .001$, trials. Since all relevant tests retain their significance when this averaging is performed, it is unlikely that our results are driven by any effects related to multiple births.

**Vineland-II: Adaptive Behavior Scales**

The Vineland-II questionnaire (see Materials and Apparatus for more details) is a survey that assesses children’s abilities in four different domains: communication, daily living, socialization, and motor skills. Additionally, scores in each domain are combined using a transformed sum to give an overall score and percentile rank for each child relative to average scores for children in the same age range (Adaptive Behavior Composite Scores).

As the Vineland-II is a standardized questionnaire, we were able to compare all preterm infant’s scores to average scores for 6 month old infants. We examined Vineland-II results for preterm infants included in the fNIRS experiment reported in the main text. Four additional infants were excluded from these analyses due to a lack of complete Vineland score information. Additionally, four more infants were excluded from analysis of SES because of missing parental education or income information.

The Adaptive Behavior Composite Scores (a transformed sum to give an overall score for each infant) were significantly lower than average, $t(47) = -4.76$, $p < .001$, placing the preterm infants we examined just above the 30th
percentile for infants of their age group. However, there is evidence of systematic parental bias in the responses to this questionnaire. In contrast to the Composite Scores, infant’s scored significantly higher than average in all Vineland sub-domains: communication, $t(47) = 25.62, p < .001$, daily living skills, $t(47) = 37.90, p < .001$, socialization, $t(47) = 19.69, p < .001$, and motor skills, $t(47) = 38.73, p < .001$. The high scores in individual domains is possibly indicative of parental reporting bias, suggesting that the combined score takes into account some such biases and may be the more robust assessment of developmental stage. Note that in the motor skills domain the distribution of scores was tightly clustered between 90-95 making scores outside this range carry undue weight in subsequent analyses. This means that any significant results reported below that involve correlation with this domain are likely due to chance.

Next, we examined particular demographic variables (many of which were explored in our above analysis of neural response data) in relation to Vineland scores predicting a positive relationship between these variables. SES significantly predicted Adaptive Behavior Composite Scores in the preterm infants, $b = -1.09, t(35) = -3.66, p < .001$, and likewise explained a significant proportion of the variance in these scores, $F(1,35) = 13.41, p < .001, R^2 = .28$ but with a negative relationship between SES and the scores. Similarly, SES also significantly predicted all sub-domains (see Table 1). These negative correlations, indicating that infants with higher SES performed worse on the Vineland assessment, is somewhat surprising given our prediction that SES is, if anything, positively correlated with developmental stage. However, this relationship could potentially be attributed to reporting bias. Notably, SS also find a negative correlation between “social class” and comprehension in parental questionnaires and similarly attribute this relationship to a reporting bias. Again, we find evidence that the Vineland-II scores were strongly affected by parental reporting biases.

We also explored the relationship between gestational age at birth and Vineland performance predicting a positive relationship. We found that gestational age was a significant predictor of Adaptive Behavior Composite Scores, $b = -0.20, t(39) = -2.27, p = .0292$ but again with a negative relationship. It also significantly predicted scores in the Daily Living sub-domain, $b = -0.31, t(39) = -3.22, p = .00256$, but did not significantly predict scores in any other sub-domain. Gestational age explained a significant proportion of variance in both Composite scores, $F(1,39) = 5.13, p = .0292, R^2 = .12$. and Daily Living scores, $F(1,39) = 10.39, p = .00256, R^2 = .21$. Given the strong relationship between gestational age at birth and SES ($r = .60, t(35) = 4.42, p < .001$), it is likely that, once again, this unexpected negative correlation could be due to bias in parents’ reporting of their children’s behaviors. This hypothesis is further supported by the fact that we see no dependence of neural response on gestational age (see Results above).

To examine the potential relationship between multiple births and Vineland results, we split preterm infants into those that were born as part of a multiple birth ($n = 22$) and those that were born as a single birth ($n = 19$) and compared their Vineland scores in all domains. There was a significant difference between
multiple and single birth infants in Composite scores, $t(36.01) = −2.42$, $p = .02075$, as well as Communication, $t(35.88) = −2.29$, $p = .02812$, and Daily Living scores, $t(30.50) = −2.88$, $p = .007235$. The scores for infants born as members of a multiple birth were significantly lower than scores for infants born in single births.

We found that sex only marginally significantly predicted Communication, $b = 8.10$, $t(39) = 1.93$, $p = .0614$, and Socialization, $b = 4.62$, $t(39) = 1.80$, $p = .0794$, and was not a significant predictor of scores in any other domains. Similarly, it explained a small, yet marginally significant proportion of the variance in scores in these two domains (Communication: $F(1,39) = 3.71$, $p = .06145$, $R^2 = .09$; Socialization: $F(1,39) = 3.25$, $p = .07935$, $R^2 = .08$). This indicates that the sex of the infant did not have a significant effect on their neural development or their parents’ reporting of it.

Keeping in mind that these Vineland scores should be interpreted with caution, we examined potential correlations between Vineland scores and neural response during visual omission trials (Because we consider both Vineland scores and neural response to be dependent variables in various analyses presented in this paper, we elected to perform Pearson’s product moment correlation test rather than a linear regression to examine the relationship between these two variables. $r$ values reported in these correlation tests map directly onto $R^2$ values in the equivalent linear regression.). We found no significant correlation between neural response in the occipital ROI and Vineland Composite score, $r = −.23$, $t(39) = −1.46$, $p = .1526$. In the temporal ROI, there was a marginally significant relationship between neural response and Composite score, $r = −.27$, $t(39) = −1.76$, $p = .08695$. We further examined potential relationships between Vineland scores in each sub-domain and visual omission response. These analyses are purely exploratory and no definite conclusions can be drawn from the results, especially keeping in mind the above analyses showing that there may be significant effects of reporting bias on the Vineland scores. We found a significant relationship between visual omission response and Vineland motor skills score in both the occipital, $r = −.31$, $t(39) = −2.04$, $p = .04804$, and temporal, $r = −.37$, $t(39) = −1.76$, $p = .01654$, ROIs, but no significant relationship in any other Vineland sub-domains. As described above, the tight clustering of Motor Skills scores makes it difficult to interpret any significant correlations with these scores.

Control Behavioral Experiment

In the main text, we report differences in neural responses to unexpected visual omissions between infants born preterm and full-term. We assert that this is a result of deficits in top-down prediction signals in preterm infants. However, an alternative explanation is that the visual omissions are less unexpected for preterm infants and result from differences in preterm and full-term infants abilities to detect the visual omission. For example, if preterm infants fail to learn the audiovisual association, the omission of a visual stimulus after a sound will not be unexpected. This alternative explanation shifts the origin of
the neural differences towards the systems necessary to detect visual omissions (e.g., differential associative learning or statistical learning abilities) rather than the top-down sensory predictions arising from the association.

This highlights the difference between the use of prediction and prediction errors in the domain of reinforcement learning and the top-down sensory predictions being studied here. Prediction supporting reinforcement learning originates in the basal ganglia and other subcortical circuitry and is distinct from the top-down prediction of sensory input that modulates the cortex that is the focus of the current study. Thus, while it is possible that prematurity affects reinforcement or associative learning early in life, the current task was designed to place relatively minimal demands on these types of learning systems, and instead to probe infant’s abilities to use recently acquired information to generate top-down predictions about their upcoming sensory input. Specifically, audiovisual stimuli are presented with substantial temporal overlap reducing any need for processes like prediction or memory for the association. There is also a familiarization period at the beginning of the experiment to help infants form the association without any violations to the structure, followed by 80% of the trials providing a consistent audiovisual presentation to help further associative learning and maintain any predictions about sensory input.

However, in addition to the careful design of this study, we can directly test whether these crucial assumptions of the experiment are met. This Control Behavioral experiment addresses these potential qualifications by directly comparing preterm and full-term infant’s abilities to detect visual omissions after the same type of exposure they would have received in the fNIRS experiment. This control experiment allows us to determine whether both groups are similarly sensitive to the kind of violation that we employed in the current study (i.e., visual omissions).

Participants

One hundred additional infants were recruited for the control study: 50 full-term infants, 50 infants born prematurely. These groups of infants were recruited using the same methods as reported in the main text with the same exclusionary criteria. Of these infants, 12 were not included in the analyses because they either didn’t watch for a sufficient amount of time (didn’t watch for at least 2 of the 4 test trials, 11 infants, 6 preterm) or due to technical issues (1 full-term infant). For the 43 preterm infants included in the final analyses, average gestational age at birth is 30 weeks and 3 days, 21 were female. For race, 16 infants were identified as black, 2 were multiple races, 2 did not report and the remainder were white. For ethnicity, 6 infants were identified as hispanic, 4 did not report their ethnicity and the remainder were identified as non-hispanic. Average age at participation was 5.7 months (corrected age).

Of these preterm infants, 12 (27.9%) were part of multiple births (twins or triplets) and the rest were singletons. Eleven (25.5%) were exposed to prenatal substance abuse (including tobacco and alcohol use) but had no known health issues resulting from this exposure. All included infants but one had normal
hearing screenings with one infant having mid-to-moderate hearing loss bilaterally. This infant was included in the analysis as their hearing deficits were not severe. Another infant was excluded due to an abnormal hearing screening indicating moderate-to-severe hearing loss bilaterally. No infants had laser eye surgery for ROP; All but 4 infants were discharged with fully mature retinas. Six (14%) infants had evidence of intraventricular hemorrhage (all Grade I, 4 bilateral, 1 unilateral left, 1 unilateral right). Seven (16.3%) infants were discharged on medications for BPD and three (7%) were re-hospitalized for respiratory illness (none of these three were discharged on medications).

For 44 full-term infants included in the final sample, 23 were female. For race, 1 was identified as Asian, 6 were identified as black, 3 were mixed race or other, 1 did not report and the remainder were white. For ethnicity, no infants were identified as hispanic and 2 did not report. Average age at participation was 6.5 months.

Socioeconomic status was calculated using the same measure as for the fNIRS study. For the full-term group it was on average 17.67 (out of 23) with a minimum of 8 and a maximum of 23; There was incomplete data for 7 of the infants. For the preterm group, average was 13.84, minimum of 5 and a maximum of 22.5 and 4 infants of missing data. Thus, as with the fNIRS experiment, there is a much wider spread of SES for the preterm group than the full-term group and there is a significant difference in SES between the populations (Wilcoxon Rank Sum test, W = 388, \( p < 0.001 \)).

Vineland scores were also collected for preterm infants in this experiment. The average scores were as follows: Communication, 96.5; Daily Living, 91.3; Socialization, 90.1; Motor Skills, 89.3; Adaptive Behavior, 90.16. The Adaptive Behavior score is at the 31st percentile; the same as the preterm group in the fNIRS experiment.

**Design and Procedure**

Our goal was to make the familiarization as similar to what infants in the fNIRS experiment experienced as possible. As with the fNIRS experiment, each sound is uniquely paired with a direction of movement for the smiley face and the ISI between successive stimuli is the same (jittered from 1-1.5 seconds). Thus, all the stimuli employed were identical with the exception of the fireworks video in the inter-block intervals (IBIs) which were removed because we were not recording the slow hemodynamic responses. Instead, the same music used during the IBIs was paired with a blinking white circle on a black screen. This was the attention-getter employed after an infant looked away for 2 seconds.

Familiarization was fixed and yoked to the amount of exposure that infants received based on their self-directed looking during the fNIRS experiment. In other words, the stimuli that infants viewed during the fNIRS experiment were now presented to a separate group of infants in this behavioral control experiment. Notably, in addition to seeing approximately the same number of audiovisual pairs, infants also saw a similar number of visual omissions during familiarization. Including a minority of visual omission trials is necessary to have familiarization for the current experiment be as similar as possible to the
experiences that infants had during the fNIRS study. The presence of visual omissions might negatively affect audiovisual learning and an infant’s expectation for the visual stimulus after hearing a sound. Moreover, it could be preterm infant learning is more fragile to the presence of even infrequent visual omissions. To this end, infants first saw 18 AV trials (9 of each AV pairing). This experience reflects the 3 initial blocks that infants received during the fNIRS experiment with no visual omissions. Then, infants were familiarized with another 72 AV pairs (4 more blocks of 18 trials each) with 20% of trials presenting a visual omission. This amount of exposure was chosen to reflect the number of trials (including visual omissions) that infants watched during the fNIRS experiment. This is a fixed familiarization sequence though infant looking was monitored by research staff. When infants looked away from the screen for 2 consecutive seconds, stimulus presentation was paused and the attention getter was played. When stimulus presentation resumed, it represented the previous stimulus as a 2 second look away time entails that the visual stimulus was not seen in the previous trial. This procedure was followed until infants viewed all familiarization stimuli.

After familiarization, infants were presented with test trials where a single AV pair was presented either with 50% of the trials presenting a visual omission or visual presentation included in each trial. We will refer to these as audiovisual vs. visual omission test trials. For the visual omission test trials, the first presentation was always a visual omission. On each subsequent trial, there was a 50% probability that the visual stimulus would be presented. There were 4 test trials: Two audiovisual test trials, one for each pair and the same for the visual omission trials. It was counterbalanced whether infants viewed audiovisual or visual omission test trials first. Then both test trials of that type were presented (order randomly determined), followed by the two test trials from the other type. In all 8 test trials were presented (a repeat of the first 4 trials was included in the original procedure) but only the first 4 trials were analyzed as a greater number of infants failed to complete the second set of test trials. Thus, all included infants (preterm and full-term) saw a range of 1-2 audiovisual trials (full-term mean: 1.93, preterm mean: 1.91) and 1-2 visual omission trials (full-term mean: 1.89, preterm mean: 1.91).

Looking times were determined by the real-time coding of research staff blind to which trials were visual omission or visual present. Specifically, as the only difference between trial types is the visual presentation of the stimuli, the staff member was able to hear the sounds being played (in order to determine if the experiment was running and whether the attention getter was being presented) and see a video feed of the infant but was not able to see the visual input the infant received. Thus, the research staff member was fully blind to the key experimental manipulations. The experiment continued until infants consistently stopped looking at the screen. Moreover, the caregiver was masked using a visor and blind to experimental condition.
Results and Discussion

Each test trial was presented as long as the infant consistently watched the screen. This resulted in a highly skewed distribution of looking times across test trials. In order to reduce the outliers within the population, test trials that were more than 2 standard deviations from the mean were excluded from subsequent analyses. Cut-off for exclusion was 89.33 seconds per test trial (based on mean looking of 23.90 secs and standard deviation of 32.71 secs). Infants were included in the subsequent analyses if they watched at least one test trial from each type (minimum of 2 test trials) after outlier rejection. The majority of infants watched all 4 test trials. All analyses were conducted on the average looking for each test trial type for each infant.

We first sought to determine whether there is an overall looking time preference for visual omission test trials. Using a paired t-test, we find significantly longer looking to audiovisual trials compared to visual omission trials ($t(86) = 2.76, p = 0.007$). Given that the looking time data are not normal, we ran this test again using log-transformed looking times and found the same result ($t(86) = 2.82, p = 0.006$). We also ran a non-parametric test (Wilcoxon signed rank test) and again found a significant difference in looking between test trial types ($V = 645, p = 0.04$). Thus, we robustly find that overall infants look longer to audiovisual trial providing clear evidence that infants are learning the association and detect a difference between a visual present and a visual omitted trial. Given that infants received a fixed familiarization (i.e., we did not employ infant controlled habituation), it is straightforward to interpret this pattern of results using the canonical Hunter and Ames model [59]: A familiarity preference for audiovisual trials is consistent with infants still forming their audiovisual association and yet achieved a level of internal representation that supports a novelty preference.

Regardless of the direction of looking-time preference, the crucial question is whether this looking-time pattern differs between preterm and full-term infants. To this end, we conducted a 2x2 ANOVA with test trial type (audiovisual/visual omission) and birth status (full-term/preterm). Again we confirm the difference in looking across test trial types with a main effect of test trial type ($F(1,170) = 5.98, p = 0.02$). We do not find a main effect of birth status nor an interaction between birth status and test trial type ($ps >0.5$). We again ran this analysis with log-transformed looking times and confirm these results.

Importantly, we sought to determine whether the difference in looking time is robust in preterm infants. Including only preterm infants in the analyses, we find significantly longer looking to audiovisual trials ($M = 23.71$ secs) compared to visual omission trials ($M = 16.54$ secs) in the paired t-test ($t(42) = 2.44, p = 0.02$), the Wilcoxon signed rank test ($V = 645, p = 0.04$) and a marginal difference with the t-test of the log-transformed looking times ($t(42) = 1.93, p = 0.06$). We find evidence of the same pattern of looking for full-term infants (Visual Omission $M = 22.73$ secs; Audiovisual $M = 18.64$ secs) but this effect only reaches marginal significance: The paired t-test and Wilcoxon are not significant ($t(43) = 1.48, p = 0.15; V = 617, p = 0.16$) but the t-test with the
log transformed data is \((t(43) = 2.02, \ p = 0.05)\). Applying the outlier cutoff to full-term infants only (cutoff = 80 seconds), results in all these tests being either significant or marginal for full-term infants suggesting that these infants do exhibit longer looking to visual omissions but that preterm and full-term infants have slightly different distributions of looking times.

In sum, we examined looking-time preferences, a canonical behavioral measure of internal representations in infancy, in order to determine whether preterm and full-term infants have similar detection of visual omissions. We hypothesized that there would be no differences in this ability across groups as the task was designed to reduce learning demands (e.g., temporal overlap). Consistent with this hypothesis, we find evidence that both groups have preferences to look at audiovisual trials over visual omission trials, and there are no differences between the groups. Thus, we find strong evidence that preterm and full-term infants similarly detect of a visual omission after the type of experience that they received in the fNIRS experiment reported in the main text. Thus, the neural differences between these groups is not readily attributable to differences in the assumptions of the tasks (e.g., learning) providing evidence that prematurity effects top-down sensory prediction abilities specifically.
Supplemental References


