Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months

Lauren L. Emberson[a,b,c,1], John E. Richards[d,e], and Richard N. Aslin[b,c,1]

[a]Psychology Department, Princeton University, Princeton, NJ 08542; [b]Brain and Cognitive Sciences Department, University of Rochester, Rochester, NY 14627; [c]Rochester Center for Brain Imaging, University of Rochester, Rochester, NY 14627; [d]Department of Psychology, University of South Carolina, Columbia, SC 29208; and [e]Institute for Mind and Brain, University of South Carolina, Columbia, SC 29201

Recent theoretical work emphasizes the role of expectation in neural processing, shifting the focus from feed-forward cortical hierarchies to models that include extensive feedback (e.g., predictive coding). Empirical support for expectation-related feedback is compelling but restricted to adult humans and nonhuman animals. Given the considerable differences in neural organization, connectivity, and efficiency between infant and adult brains, it is a crucial yet open question whether expectation-related feedback is an inherent property of the cortex (i.e., operational early in development) or whether expectation-related feedback develops with extensive experience and neural maturation. To determine whether infants' expectations about future sensory input modulate their sensory cortices without the confounds of stimulus novelty or repetition suppression, we used a cross-modal (audiovisual) omission paradigm and used functional near-infrared spectroscopy (fNIRS) to record hemodynamic responses in the infant cortex. We show that the occipital cortex of 6-month-old infants exhibits the signature of expectation-based feedback. Crucially, we found that this region does not respond to auditory stimuli if they are not predictive of a visual event. Overall, these findings suggest that the young infant's brain is already capable of some rudimentary form of expectation-based feedback.

perceptual development | associative learning | fNIRS | infant | multisensory

Over the past two decades, theoretical focus has shifted from predominantly feed-forward hierarchies of cortical function, where sensory cortex propagates information to higher level analyzers on the way to decision and motor control areas, to models in which feedback connections to lower-level cortical regions allow extensive top-down functional modulation based on expectation (1). An influential model for incorporating feedback is predictive coding, which compares expectations or predictions to input at each level of the processing hierarchy (2–4). There is extensive and compelling evidence for expectation-based modulation even at the earliest levels of sensory processing in both humans (5–7) and nonhuman animals (8). Because complex, naturalistic sensory input is characterized by temporal, spatial, and contextual regularities, the ability to modulate early sensory function as a result of expectations is believed to support adaptive perceptual abilities (4).

Empirical evidence for expectation-based feedback, however, is restricted to adult humans and nonhuman animals. With their extensive experience, adults already have developed sophisticated internal models of the environment and have a highly interconnected brain and efficient neural processing. Comparatively, human infants are born with a demonstratively immature behavioral repertoire, have underdeveloped sensorimotor and cognitive capacities, and lack sophisticated internal models of the environment. These internal models undergo substantial postnatal development even in the early sensory cortex. For example, spontaneous activity of primary visual cortex (V1) in

ferrets converges with evoked visual responses to naturalistic stimuli across development but does not converge with evoked responses to nonnaturalistic stimuli (9). Thus, statistics in sensory input progressively create an adaptive, internal model of the environment that is evident in V1, and, therefore, the neonate brain does not have access to sophisticated expectations about sensory input. Moreover, converging evidence from both anatomical (10) and functional connectivity analyses (11) has demonstrated that human infants have a much less interconnected brain and instead have a predominance of local connections and a slow development of long-range interactions (12). Finally, human infants exhibit characteristically slower neural processing even for basic sensory stimuli (13).

Given the considerable differences between infant and adult brains, it is unclear whether the infant brain is capable of the sophisticated expectation-based modulation of sensory cortices that has been observed in adult brains. Relatedly, the developmental process by which feedback comes to modulate neural function based on changing expectations is entirely unknown. One possibility is that infants do not exhibit expectation-based feedback, but rely on a predominantly feed-forward architecture. In this case, a feed-forward neural architecture would be a developmental precursor to the feedback mechanisms that would emerge after extensive postnatal experience and would be consistent with the immaturity of infant cognition and sensorimotor control. Another possibility is that the fundamental

---

Significance

Although infants are excellent learners, it is unclear whether infants use neural strategies similar to those of adults to track changes in their environment. One adult neural strategy is to use feedback connections to modulate sensory cortices based on their expectations. The current study provides, to our knowledge, the first evidence that the fundamental architecture required for sensory feedback is already in place in infancy. This top-down modulation is especially impressive because the study employs an audiovisual task that requires the flexible use of long-range neural connections, and the infant brain is dominated by short-range neural connections with weak (e.g., unmyelinated) long-range connections. These results suggest that learners can use sophisticated top-down feedback neural strategies from an early age.

NEUROSCIENCE

PSYCHOLOGICAL AND COGNITIVE SCIENCES

Fig. 1. Schematic of the two trial types: A+V+ (audiovisual trials) and A+V− (visual omission trials). After a period of familiarization to only A+V+ trials, infants saw visual stimuli accompanying auditory stimuli 80% of the time. Thus, infants expected to see the visual stimulus after each auditory stimulus.

architecture that allows expectation-based feedback is present in early infancy and provides a scaffold for the development of efficient internal models of the environment. Here, we ask whether expectation-based feedback is evident in the infant brain: specifically, in the sensory cortex.

Despite extensive behavioral evidence that infants can form expectations in many perceptual and cognitive domains (14, 15) and are able to quickly and robustly respond to new information in their environment (e.g., statistical learning abilities) (16, 17), it remains unclear how their expectations about future sensory input are instantiated neurally. The infant event-related potentials (ERPs) literature provides a glimpse but falls short of clarifying the underlying neural mechanism. The traditional oddball design, which presents a frequent stimulus and an infrequent stimulus (e.g., 80% and 20% respectively), reveals a robust novelty response in several waveform components (delayed version of the P300 and the mismatch negativity) (18–20). However, the ERP is exquisitely sensitive to the properties of the stimuli, and so the enhanced response to the infrequent stimulus could be a prediction error, a nonspecific surprise effect, or a recovery to the greater repetition of the frequent stimulus (i.e., rebound from repetition suppression) (21). Using similar designs with similar interpretative limitations, recent work with functional near-infrared spectroscopy (fNIRS) (an optical imaging method for noninvasively recording functional changes in the hemodynamics of the infant cortex while the infant is awake and behaving) (22, 23) has revealed increased neural responses to the presentation of novel acoustic stimuli (24) and evidence of stimulus anticipation (25).

A more interpretively transparent design for assessing whether expectations can modulate responses in sensory cortex is the stimulus-omission paradigm. Here, two or more stimuli are presented in a predictable temporal order, and one of the expected stimuli is occasionally and unexpectedly omitted. Because there is no sensory input during stimulus omission, the response cannot be due to recovery from repetition suppression, and, if the effect is localized (i.e., does not occur in all cortical regions), it cannot be due to nonspecific surprise. This stimulus-omission paradigm has been used with mice to record from V1 with multielectrode arrays (8), with human adults to record from auditory and frontal areas with magnetoencephalography (MEG) (7), from V1 using fMRI (26), and with presurgical epilepsy patients to record from temporal-parietal areas with cortical electrodes (27). A particularly impressive variation is the ability of a recently learned cross-modal

association (e.g., audiovisual stimuli) to generate responses to the unexpected omission of one of the previously paired stimuli, as recently demonstrated in human adults (5, 28).

We used the cross-modal stimulus omission paradigm to determine whether a recently learned audiovisual association drives occipital cortex responses during an unexpected visual omission in young human infants. We presented 36 6-month-old infants (*Materials and Methods*) with novel sounds and visual stimuli such that a sound predicted the presentation of a visual stimulus. Infants can rapidly learn arbitrary audiovisual associations even from a young age (29, 30). Therefore, after a brief period of familiarization (less than 2 min, 18 audiovisual or A+V+ events), infants were presented with trials where the predictive auditory stimulus was presented and the visual stimulus was unexpectedly withheld (see Fig. 1 and *Materials and Methods*). These unexpected visual omission trials violate sensory expectations but do not present any sensory input. Therefore, any occipital cortex responses to unexpected visual omissions were evidence of a neural response generated as a result of a violation of sensory expectation and not a result of lower-level neural adaptation effects, such as repetition suppression or generalized novelty responses.

Results

To calculate the magnitude of the hemodynamic response, normalized changes in blood oxygenation were averaged from 5 to 9 s after stimulus onset within two neuroanatomically defined regions of interest (ROIs) (occipital, three NIRS channels; temporal, five NIRS channels) (Fig. 2) [see *Materials and Methods* for details on the magnetic resonance (MR)–fNIRS coregistration method (31) and selection of our analysis window]. We considered changes in oxygenation for trials where a predictive auditory stimulus was presented but the visual stimulus unexpectedly was omitted (unexpected visual omissions or A+V− trials) and to trials where both an auditory and visual stimulus were presented (A+V+ trials, consistent with overall exposure) (Fig. 1). Although, overall, the auditory and visual stimuli were presented together 80% of the time (after a period of familiarization), our experimental design allowed us to compare equal numbers of A+V+ trials and A+V− trials (*Materials and Methods*).
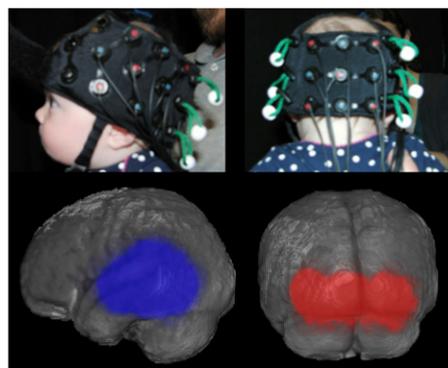
Fig. 2. (*Top*) Pictures were taken of each infant who participated in the study to help determine the location of the NIRS optodes in relation to anatomical markers. (*Bottom*) Two regions of interest (ROIs) were temporal (comprised of five channels) and occipital (comprised of three channels). ROIs were selected based on their neuroanatomical localization for the majority of infants. All channels within each ROI were averaged together. Three average MR templates were used for the coregistration to capture natural variability in infants' head sizes in the age range studied (*Materials and Methods*); this figure presents the range and probability (higher probability is indicated by brighter colors) of NIRS recordings for the two ROIs as displayed on the MR template used for coregistration of all of the infants included in the final sample.
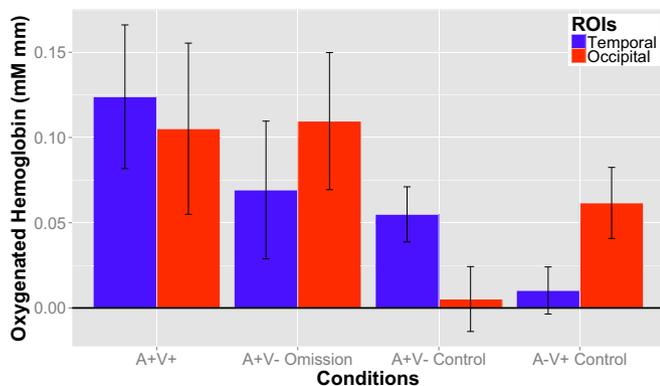
Fig. 3. Average change in oxygenated hemoglobin in two ROIs (temporal and occipital cortex) in response to (going left to right) A+V+ trials that were consistent with 80% of stimuli presented to infants after familiarization, A+V− omission trials, or unexpected visual omissions where the auditory stimulus was presented but the visual stimulus was not. These trials occurred 20% of the time. A separate group of infants was presented with two types of control trials: expected visual omissions (A+V− control) and expected auditory omissions (A−V+ control). The level of oxygenated hemoglobin was averaged from 5 s after stimulus onset to the mean duration of the jittered ISI, which was 9 s post-stimulus onset.

First, considering the A+V+ trials, we found significant increases in oxygenation from baseline in both the temporal and occipital ROIs [$ts(16) > 2.44$, $Ps < 0.026$] (Fig. 3). This important confirmatory result is consistent with previous work on temporal responses to auditory stimuli and occipital responses to visual stimuli in infants (see ref. 22 for a recent review).

The Infant Occipital Cortex Responds During Unexpected Visual Omissions. Next, we consider responses in the same ROIs to trials where a predictive auditory stimulus was followed by an unexpected omission of the visual stimulus. Crucially, we found a robust occipital cortex response [$t(16) = 4.81$, $P = 0.00019$] (Fig. 3) that was not statistically distinct from a response to the presentation of a visual stimulus [A+V+ trials compared with A+V− trials, $t(16) = −0.10$, $P = 0.92$]. Thus, in an ROI that responds robustly to the presentation of a visual stimulus, we also found a robust response during the unexpected omission of a visual stimulus. We also found significant activity in the temporal cortex in A+V− trials [$t(16) = 2.64$, $P = 0.018$]. However, this activation was significantly reduced from the A+V+ trials [$t(16) = 2.58$, $P = 0.020$] (Fig. S1). Similar to previous work (32), we found that audiovisual information increased occipital lobe activity compared with visual only presentation. Future work is needed to carefully disentangle whether it is simply the superposition of auditory and visual information that results in augmented sensory cortex activity or whether additional predictive processing or learning is contributing to this sensory cortex modulation.

The Occipital Cortex Does Not Respond During Expected Visual Omissions. In an additional experiment, we conducted a crucial control: How does the infant occipital cortex respond when it is presented with expected visual omissions? A separate group of infants with comparable demographics (*Materials and Methods*) were presented with the same auditory and visual stimuli, but now these stimuli are always presented separately: The auditory stimulus was never paired with the visual stimulus, and the visual stimulus was never presented with the auditory stimulus. Thus, when infants heard the auditory stimulus, they did not expect to see the visual stimulus: In the absence of audiovisual association, visual omissions are now expected. Identical fNIRS recording and analysis methods revealed that expected visual omissions

(A+V− control trials) did not result in a response in the occipital cortex [$t(18) = 0.34$, $P = 0.74$] (Fig. 3).

In a direct comparison across experiments, we found that the occipital cortex responded differentially to the same stimuli (the presentation of an auditory stimulus without a visual stimulus, A+V−) depending on the presence or the absence of visual expectations [$t(34) = 3.35$, $P = 0.002$] (Fig. 3, center two columns). Fig. 4 presents the time course of responses in the occipital cortex in response to the same stimuli but in the presence or absence of a visual expectation. When infants expected visual input but none was received, there was an increase in cortical activity in this region.

In this separate group of infants, we also confirmed that the same occipital ROI robustly responds to the presentation of a visual stimulus (A−V+ control trials) (Fig. 3, rightmost column, $t(18) = 3.8537$, $P = 0.0012$). Although numerically the responses to both A+V+ trials and unexpected visual omissions (A+V− trials) were higher than for the control visual only trials (A−V+), this difference was not statistically significant [$ts(34) < 1.53$, $Ps > 0.13$] (Fig. S2). However, we found a significant increase in temporal cortex response to audiovisual A+V+ trials compared with unexpected visual omission trials (A+V−, reported earlier, Fig. S1) and a marginally significant increase compared with an expected visual omission [control A+V−, $t(33) = 2.00$, $P = 0.053$], suggesting that the presence of concurrent audiovisual information modulates activity in the temporal cortex, with some suggestion for the same in the occipital cortex.

Discussion

Although decades of research have used violations of expectation to probe infant behavior (14, 15), the current study, to our knowledge, is the first to determine that the infant occipital cortex responds to a mismatch between an infant's visual expectation and current sensory input. Trials that violated visual expectations did so by omitting the presentation of a visual stimulus, and therefore no novel stimulus was presented. This design is crucial because it circumvents the possibility that these results are driven by lower-level adaption responses, such as repetition suppression. The inclusion of MRI coregistration of fNIRS recordings with age and head size-appropriate MRI templates provided clear neuroanatomical constraints and interpretations on our functional data. Even though we used broadly defined occipital and temporal ROIs, MR coregistration was necessary to determine which probes on the back of the head were recording from the occipital cortex vs. the cerebellum and
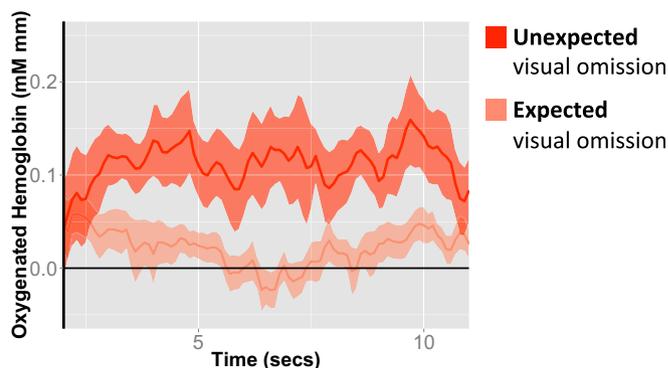
Fig. 4. Time course of changes in oxygenation for the same stimuli (A+V−) but when a visual omission was expected and unexpected. Time is in seconds from the onset of the auditory stimulus (visual stimulus started 750 ms after auditory stimulus). These time courses correspond to the mean changes in oxygenation for the occipital ROI as represented in the two middle columns of the bar plot in Fig. 3. In all plots, error bars represent the SEM.

which probes on the side of the head were recording from the temporal vs. the frontal lobe. Without conducing MR coregistration, we would have mistakenly localized probes to the wrong lobe (e.g., frontal channels mistaken for temporal). In addition, we now know which subregions of the occipital and temporal lobes were recorded from during the study, paving the way for future work to refine current analyses to subregions of these perceptual cortices. Broadly, we found that the same ROI that robustly responds to visual stimuli (both within the same group of infants for A+V+ trials and in a separate group of infants in the control experiment for A−V+ trials) exhibits a robust response when an infant expects to see a visual stimulus, based on an auditory cue, but this visual stimulus is withheld. Crucially, we found that this region does not respond to auditory stimuli if they are not predictive of a visual event.

Overall, these findings suggest that the young infant's brain is already capable of some rudimentary form of expectation-based feedback across anatomically distinct regions. The adult brain has a highly interconnected, sophisticated architecture that readily exhibits changes based on expectations (2, 3, 33). Young infants presumably lack both the sophisticated internal models of the environment and the well-honed sensorimotor processing abilities of adults and are slower to process even basic stimuli (13). In addition, connections in the infant brain are not mature and experience a robust developmental trajectory. Research on the development of cortico-cortical interactions in nonhuman animals has found that, within the occipital lobe, there exist both feedback and feedforward connections early in development, but, whereas some aspects of a visual hierarchy exist early, there is a dominance of feedforward connections. With visual experience, the balance of these types of connections shifts with both an increase in feedback connection and a decrease in feedforward connections (34). fMRI studies of infant functional connectivity using resting state measures also show an abundance of short-range connections, but, critically for the current study, a paucity of long-range cortical connections between anatomically distinction regions (10, 11). Despite these differences between adults and infants, we still find that infants do not rely on exclusively feedforward signals for evaluating their sensory input in relation to their expectations. This finding differs from classic models that emphasize the role of bottom-up, feedforward projections (35) and often argue for the encapsulation of visual information from top-down knowledge or expectations (36). Instead, we found that the infant brain is capable of exhibiting rapid, flexible top-down influence on lower level sensory regions as a result of a cross-modal expectation. Whether the rudimentary expectation-based feedback observed here in the infant brain is indicative of predictive coding—where only an error signal is passed along to higher levels of the visual hierarchy—will require future work and may be beyond the scope of current developmental neuroscientific methods. Importantly, the current task requires the rapid recruitment of long-range connections between the temporal and the occipital lobe, which is beyond the scope of the early feedback connections found within the visual system (34). One possibility, supported by adult neuroimaging with a similar task (5), is that the expectation-based modulation exhibited in the current study is not mediated by cortico-cortical connections but is established subcortically via learning and memory systems perhaps within the basal ganglia.

Despite this evidence for early availability of top-down signals in the occipital cortex, it is clear that neural feedback systems are not mature at 6 months of age. Behavioral evidence suggests a relatively protracted trajectory of developmental improvement on visual tasks that require extensive feedback, including visual cue combination (37) and figure-ground segregation (38). If rapidly learned top-down influences are already present in the infant brain, what other factors are responsible for the poorer neural processing and cognitive capacity in infancy and the more protracted

development of visual skills that require feedback? For example, perhaps the poor internal models prevent accurate or well-defined predictions. Thus, these models may be immature both in terms of the perceptual representations and/or the higher level systems that might integrate information to help generate predictions.

Although rarely discussed directly, learning from experience with structured input is the essential component in demonstrations of expectation-based feedback, regardless of the age of the subject (for example, learning the sequence of presentation of visual stimuli) (8). Indeed, the current paradigm can be considered in the context of classical conditioning, where a predictive stimulus (the auditory event) conditions the occipital cortex through predicting a visual event. The experience that infants will use to develop their perceptual, cognitive, and motor abilities is similarly structured: our daily sensory input reflects the structured environment in which we live (4, 39–41), and therefore it is not randomized or counterbalanced like the majority of studies that have been used to probe neural function. Our sensory input is structured along many spatiotemporal dimensions. This structure permits expectations, predictions, stimulus templates, and/or internal models to be built when paired with a neural architecture, which permits flexible and rapid learning or adaptation to these new stimulus statistics. Despite these clear intersections, it is not clear empirically or theoretically how prediction and learning interrelate. Is prediction part of the learning process? Is prediction the outcome of learning, where an acquired representation of the environment enabled prediction of upcoming sensory input? Answers to these questions will require future research.

More broadly, the elucidation of expectation-based responses in the infant occipital cortex creates a line of continuity between young infants and adults. The brain's ability to adaptively shape cortical activity to the structure of the environment may rely on the same underlying neural architecture from early infancy into adulthood. Despite the growth and maturation of the brain during development, with all of its functional and anatomical changes, some of the basic neural mechanisms may be invariant. The present study raises the possibility that these adaptive neural mechanisms are one of the invariant mechanisms and can support life-long flexibility, plasticity, and learning. Although cortical representations and internal models of the world certainly vary widely across development, an architecture that generates experience-based expectations and propagates those expectations to early sensory regions of the brain could be an evolutionarily efficient scheme for supporting development and cognitive flexibility in a changing world.

Materials and Methods

Participants. This paper reports two experiments: (*i*) the unexpected visual omission experiment, where auditory stimuli predicted visual stimuli by preceding them temporally and infants were then presented with unexpected visual omissions and (*ii*) the control experiment, where infants were presented with the same stimuli but always in unimodal single trials (i.e., they were never paired). Infants, aged 5–7 mo, were recruited from the same population (the database of interested participants for the Rochester Baby Lab). Experiments were approved by the Institutional Review Board of the University of Rochester and informed consent was obtained from the caregivers of the infants. Infants were selected to have been born no more than 3 wk before their due date and to have had no major health problems or surgeries, no history of ear infections, or any known hearing or vision difficulties. Out of 51 infants recruited (unexpected visual omission, 26 infants; control, 25 infants), 15 infants (29% of the total sample) were excluded for failing to watch the video to criterion (8 infants), poor optical contact [6 infants (e.g., due to too much dark hair); see *fNIRS Data Analysis* for specific definitions for each of these exclusionary criteria], and experimenter error (1 infant). The final sample was 36 infants (unexpected visual omission, 17 infants; control, 19 infants; mean age, 5.7 mo; SD = 0.59 mo; 22 female; race, 31 white, 5 other or more than one, 2 Hispanic, 34 non-Hispanic).

Stimulus Presentation and Experimental Procedure. The same basic auditory and visual stimuli were used in all experiments, but in different configurations.

Stimuli appear in Fig. S3, Movie S1, and Audio Files S1 and S2. All trials started with the presentation of a monochromatic gray screen with a white box (black bordered) presented in the middle. The box was 15.2 degrees of visual angle squared (11.5cm^2 with the infant sitting ~43 cm from the screen). Immediately after, a combination of auditory and/or visual stimuli were presented: two novel, nonspeech auditory stimuli (well-described as a honk, like from a clown horn, and an unusual rattle sound) and/or two visual stimuli that are a red cartoon smiley face that was presented in two different events or ways (entering a white box from either the top or the bottom, moving into the box to touch the opposite side of the box in 500 ms, and then exiting the box in the same side that it entered from in another 500 ms). Both the auditory stimuli and the duration of the presence of the red smiley face were 1 s in duration. In the visual omission experiment, an auditory stimulus was presented 750 ms before the visual stimulus. Thus, there was overlap between these stimuli for 250 ms. An individual sound (one of two) was consistently paired with one direction of movements for the visual stimulus, creating two pairs of audio-visual stimuli which was done to balance the constraints of keeping infants sufficiently engaged to watch the requisite number of trials and having robust associative learning for the infants. Individual pairs were used with the same frequency throughout the experiment. In the control experiment, the same four stimuli (two each of auditory and visual) were presented, but always in unimodal single trials. After the conclusion of the last stimulus, the empty white box and the gray screen (static visual presentation with silence) was presented for 1,000–1,500 ms (randomly determined). 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 67 dB. Stimuli were presented using MATLAB for Mac (R2007b) and Psychtoolbox (3.0.8 Beta, SVN revision 1245).

There were two types of stimulus presentation: blocks and single trials. Blocks, presented only for the unexpected visual omission experiment, consisted of six A+V+ trials: three of each specific pairing, randomly ordered and each separated by the jittered interstimulus interval (ISI) (1–1.5 s). Single trials were used in all experiments and consisted of the presentation of a single A and/or V stimulus, followed by the same jittered ISI. In between single trials and blocks, baseline stimuli were presented [dimmed fireworks video (42) and calming instrumental version of "Camptown Races," (*Baby Music*, album released 2010)] for a jittered ISI of 4–9 s (mean, 6.5 s) (43).

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 unexpected visual omission experiment began with three blocks of A+V+ presentation (18 total pairings) to provide infants with some initial learning of the specific AV pairs and generate appropriate sensory expectations.

After familiarization, infants were presented with three types of trials, each followed by baseline: one block of six A+V+ pairs, two single trials of A+V+ pairs, and two single omission trials or A+V− which results in one block and four single trials that were presented repeatedly in a fully randomized order until the infant became fussy or consistently inattentive. This design (*i*) ensured that only 20% of the trials presented unexpected visual omissions to maintain sensory expectations for pairings over the duration of the experiment and yet (*ii*) provided equal numbers of trials and trial types to compare across A+V+ trials and omission trials (both single trials, presented with the same frequency and at consistent times throughout the experiment). For the control experiment, no familiarization was used, again based on the assumption that infants do not start the experiment with such AV expectations. Instead, only single unimodal trials were presented (two of each auditory and visual stimulus, randomly ordered in groups of four) until the infant became fussy or inattentive.

The experiment was conducted in a darkened room with dark floor-to-ceiling curtains surrounding the infant and the caregiver. Only the monitor (Tobii eye tracker) was visible to the infant because all other equipment (e.g., speakers, computers) were on the other side of the curtains and out of sight. Infants sat on their caregivers' laps. Caregivers were instructed to not interfere with the infant's watching of the video but to make sure that infants did not grab at the cap on their head (see next section) or rub up against them with the cap to move it. We also asked that they encourage the infant to be as still as possible but to 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 underneath the monitor.

fNIRS Recording. fNIRS recordings were conducted using a Hitachi ETG-4000 with 24 possible NIRS 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 3 × 3 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 NIRS channels were most likely to record from temporal and occipital cortex in infants. 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 (seven for the lateral pad over the ear and five 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 the MATLAB on the stimulus presentation computer to the Hitachi ETG-4000 using standard methods. 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, and single omission trials). The raw data were exported from the Hitachi ETG-4000 to MATLAB (version 2006a for PC) for subsequent analyses with HomER 1 (Hemodynamic Evoked Response NIRS data analysis GUI, version 4.0.0) using the default preprocessing pipeline of the NIRS data. First, the raw intensity data were normalized to provide a relative (percent) change by dividing the mean of the data (HomER 1.0 manual), thus any change from zero is meaningful and does not require an explicit baseline period. Then, the data were low-pass and high-pass filtered (two separate steps) to remove high frequency noise such as Mayer waves and low frequency noise such as changes in blood pressure. Second, changes in optical density were calculated for each wavelength, and a PCA analysis was used to remove motion artifacts. Finally, the modified Beer–Lambert law was used to determine the changes (delta) in concentration of oxygenated and deoxygenated hemoglobin for each channel (the DOT.data.dConc output variable was used for subsequent analyses, see the HomER Users Guide for full details) (44). 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 R2007b.

fNIRS Data Analysis. Subsequent analyses were conducted in MATLAB (R2013a) with custom analysis scripts. These analyses consisted of removing any additional motion artifacts. Following Lloyd-Fox et al. (2009) (45), an algorithm was written such that, for each trial, motion was detected when the concentration (either oxygenated or deoxygenated hemoglobin) was greater or less than ± 5 mM·mm. The algorithm, then, searches both forward and backward from the portion of the signal beyond this threshold for a point where the beginning of the motion occurred. Specifically, the algorithm searched for a point in the signal that fits any of these three criteria (in order of preference): (*i*) the signal changed sign; (*ii*) when there had been 15 consecutive points where the slope has been less than 0.5 (indicating that the rapid change in the signal had ceased); or (*iii*) if neither of these criteria were met, when there had been more than 200 points since the signal was beyond the ± 5 mM·mm threshold. The benefits of this algorithm are that it does not rely on potentially biased and nonreplicable "hand coding" of motion, but that it not only determines when motion was likely to have occurred (signal values beyond ± 5mM·mm) but also identifies the places where motion likely started (the steep rise or fall of the signal). An appropriate analogy is that you might use the peaks to identify mountains, but, if you want to remove them, you need to also find out where their bases are. Once the segments of signal that were likely contaminated by motion were identified, they were removed by zeroing the signal, which is an appropriate method because the signal had been normalized in homER and had the benefit of maintaining timing and "complete" data collection during the trial. This method was applied to all infants included in the study, resulting in an average of 0.36% of the data being excluded due to motion (SD = 0.60%; 21 infants had no data excluded; the maximum excluded for a single infant was 1.88%).

Then, the continuous data were 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 plus 6.5 s). The number of complete trials was determined for each trial type, and it was evaluated whether the infant met the inclusion criteria of watching a minimum of four single trials of both types (e.g., four A+V+ trials and four visual omission trials; see *Participants* for the number of infants excluded for not watching a sufficient number of trials for each study type). After exclusion, infants looked on average for 6.7 single trials of each type (SD = 2.36) and 6.29 blocks (SD = 0.99) in the visual omission experiment and 6.9 trials of each type (SD = 1.95) in the control experiment.

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 but also 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 used. These infants were not

excluded at the point where meaningful average information could be seen to minimize experimenter bias to include or exclude participants who confirmed or denied experimental hypotheses. Moreover, the decision to include or exclude infants was made once and before group averages were determined and was not revisited. Then, the average and variance of responses for oxygenated and deoxygenated hemoglobin was determined within each ROI for each infant. A single analysis time window, 5–9 s 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 (46) and to be conservative with regard to any residual response to the previous trial

(5 s after stimulus onset is an average of 14 s 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 (43).

1. Gilbert CD, Li W (2013) Top-down influences on visual processing. *Nat Rev Neurosci* 14(5):350–363.
2. Friston K (2005) A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 360(1456):815–836.
3. Rao RPN, Ballard DH (1999) Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 2(1):79–87.
4. Summerfield C, de Lange FP (2014) Expectation in perceptual decision making: Neural and computational mechanisms. *Nat Rev Neurosci* 15(11):745–756.
5. den Ouden HEM, Friston KJ, Daw ND, McIntosh AR, Stephan KE (2009) A dual role for prediction error in associative learning. *Cereb Cortex* 19(5):1175–1185.
6. Kok P, Jehee JFM, de Lange FP (2012) Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron* 75(2):265–270.
7. Wacongne C, et al. (2011) Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proc Natl Acad Sci USA* 108(51):20754–20759.
8. Gavornik JP, Bear MF (2014) Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nat Neurosci* 17(5):732–737.
9. Berkes P, Orbán G, Lengyel M, Fiser J (2011) Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* 331(6013):83–87.
10. Nagy Z, Westerberg H, Klingberg T (2004) Maturation of white matter is associated with the development of cognitive functions during childhood. *J Cogn Neurosci* 16(7):1227–1233.
11. Smyser CD, et al. (2010) Longitudinal analysis of neural network development in preterm infants. *Cereb Cortex* 20(12):2852–2862.
12. Gao W, et al. (2011) Temporal and spatial evolution of brain network topology during the first two years of life. *PLoS ONE* 6(9):e25278.
13. Lee J, Birtles D, Wattam-Bell J, Atkinson J, Braddick O (2012) Latency measures of pattern-reversal VEP in adults and infants: different information from transient P1 response and steady-state phase. *Invest Ophthalmol Vis Sci* 53(3):1306–1314.
14. Aslin RN (2007) What's in a look? *Dev Sci* 10(1):48–53.
15. Haith MM (1980) *Rules That Babies Look By: The Organization of Newborn Visual Activity* (Psychology Press, London).
16. Kirkham NZ, Slemmer JA, Johnson SP (2002) Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition* 83(2):B35–B42.
17. Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. *Science* 274(5294):1926–1928.
18. Ackles PK, Cook KG (1998) Stimulus probability and event-related potentials of the brain in 6-month-old human infants: A parametric study. *Int J Psychophysiol* 29(2):115–143.
19. Friederici AD, Friedrich M, Weber C (2002) Neural manifestation of cognitive and precognitive mismatch detection in early infancy. *Neuroreport* 13(10):1251–1254.
20. Nelson CA, Collins PF (1991) Event-related potential and looking-time analysis of infants' responses to familiar and novel events: Implications for visual recognition memory. *Dev Psychol* 27(1):50–58.
21. Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: Neural models of stimulus-specific effects. *Trends Cogn Sci* 10(1):14–23.
22. Aslin RN, Shukla M, Emberson LL (2015) Hemodynamic correlates of cognition in human infants. *Annu Rev Psychol* 66:349–379.
23. Gervain J, et al. (2011) Near-infrared spectroscopy: A report from the McDonnell infant methodology consortium. *Dev Cogn Neurosci* 1(1):22–46.
24. Nakano T, Watanabe H, Homae F, Taga G (2009) Prefrontal cortical involvement in young infants' analysis of novelty. *Cereb Cortex* 19(2):455–463.
25. Nakano T, Homae F, Watanabe H, Taga G (2008) Anticipatory cortical activation precedes auditory events in sleeping infants. *PLoS ONE* 3(12):e3912.
26. Kok P, Failing MF, de Lange FP (2014) Prior expectations evoke stimulus templates in the primary visual cortex. *J Cogn Neurosci* 26(7):1546–1554.

27. Hughes HC, et al. (2001) Responses of human auditory association cortex to the omission of an expected acoustic event. *Neuroimage* 13(6 Pt 1):1073–1089.
28. Zangenehpour S, Zatorre RJ (2010) Crossmodal recruitment of primary visual cortex following brief exposure to bimodal audiovisual stimuli. *Neuropsychologia* 48(2):591–600.
29. Fenwick KD, Morrongiello BA (1998) Spatial co-location and infants' learning of auditory-visual associations. *Infant Behav Dev* 21(4):745–759.
30. Gogate LJ, Bahrick LE (1998) Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *J Exp Child Psychol* 69(2):133–149.
31. Lloyd-Fox S, et al. (2014) Coregistering functional near-infrared spectroscopy with underlying cortical areas in infants. *Neurophotonics* 1(2):025006.
32. Watanabe H, et al. (2013) Effect of auditory input on activations in infant diverse cortical regions during audiovisual processing. *Hum Brain Mapp* 34(3):543–565.
33. Clark A (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci* 36(3):181–204.
34. Price DJ, et al. (2006) The development of cortical connections. *Eur J Neurosci* 23(4):910–920.
35. Marr D (1982) *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (Freeman, New York).
36. Pylyshyn Z (1999) Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behav Brain Sci* 22(3):341–365, discussion 366–423.
37. Nardini M, Bedford R, Mareschal D (2010) Fusion of visual cues is not mandatory in children. *Proc Natl Acad Sci USA* 107(39):17041–17046.
38. Sireteanu R, Rieth C (1992) Texture segregation in infants and children. *Behav Brain Res* 49(1):133–139.
39. Gibson EJ (1969) *Principals of Perceptual Learning and Development* (Appleton-Century-Crofts, New York).
40. Gibson JJ (1979) *The Ecological Approach to Perception* (Houghton Mifflin, Boston).
41. Goldstein MH, et al. (2010) General cognitive principles for learning structure in time and space. *Trends Cogn Sci* 14(6):249–258.
42. Watanabe H, Homae F, Nakano T, Taga G (2008) Functional activation in diverse regions of the developing brain of human infants. *Neuroimage* 43(2):346–357.
43. Plichta MM, Heinzel S, Ehlis AC, Pauli P, Fallgatter AJ (2007) Model-based analysis of rapid event-related functional near-infrared spectroscopy (NIRS) data: A parametric validation study. *Neuroimage* 35(2):625–634.
44. Huppert TJ, Diamond SG, Franceschini MA, Boas DA (2009) HomER: A review of time-series analysis methods for near-infrared spectroscopy of the brain. *Appl Opt* 48(10):D280–D298.
45. Lloyd-Fox S, et al. (2009) Social perception in infancy: A near infrared spectroscopy study. *Child Dev* 80(4):986–999.
46. Taga G, Asakawa K (2007) Selectivity and localization of cortical response to auditory and visual stimulation in awake infants aged 2 to 4 months. *Neuroimage* 36(4):1246–1252.
47. Sanchez CE, Richards JE, Almli CR (2012) Neurodevelopmental MRI brain templates for children from 2 weeks to 4 years of age. *Dev Psychobiol* 54(1):77–91.
48. Richards JE, Xie W (2015) *Advances in Child Development and Behavior*, ed Benson JB (Elsevier, Philadelphia), Vol 48, pp. 1–52.
49. Hammers A, et al. (2003) Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp* 19(4):224–247.
50. Heckemann RA, Hajnal JV, Aljabar P, Rueckert D, Hammers A (2006) Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *Neuroimage* 33(1):115–126.
51. Shattuck DW, et al. (2008) Construction of a 3D probabilistic atlas of human cortical structures. *Neuroimage* 39(3):1064–1080.