Journal of Experimental Child Psychology 242 (2024) 105889

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Infants' top-down perceptual modulation is specific to own-race faces

Naiqi G. Xiao [a,b,*], Hila Ghersin [b], Natasha D. Dombrowski [b], Alexandra M. Boldin [b], Lauren L. Emberson [b,c]

[a] Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Ontario L8S 4L8, Canada
[b] Department of Psychology, Princeton University, Princeton, NJ 08540, USA
[c] Department of Psychology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada

A R T I C L E   I N F O

Article history:
Received 9 August 2023
Revised 27 January 2024

Keywords:
Top-down perception
Infants
Other-race effect
Prediction
Perception
Eye-tracking

A B S T R A C T

Recent studies have revealed the influence of higher-level cognitive systems in modulating perceptual processing (top-down perceptual modulation) in infancy. However, more research is needed to understand how top-down processes in infant perception contribute to early perceptual development. To this end, this study examined infants' top-down perception of own- and other-race faces to reveal whether top-down modulation is linked to the emergence of perceptual specialization. Infants first learned an association between a sound and faces, with the race of the faces manipulated between groups (own race vs. other race). We then tested infants' face perception across various levels of perceptual difficulty (manipulated by presentation duration) and indexed top-down perception by the change in perception when infants heard the sound previously associated with the face (predictive sound) versus an irrelevant sound. Infants exhibited top-down face perception for own-race faces (Experiment 1). However, we present new evidence that infants did not show evidence of top-down modulation for other-race faces (Experiment 2), suggesting an experience-based specificity of this capacity with more effective top-down modulation in familiar perceptual contexts. In addition, we ruled out the possibility that this face race effect was due to differences in infants' associative learning of the sound and faces between the two groups. This work has important implications

* Corresponding author at: Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Ontario L8S 4L8, Canada.

E-mail address: xiaon8@mcmaster.ca (N.G. Xiao).

https://doi.org/10.1016/j.jecp.2024.105889
0022-0965/© 2024 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by/4.0/).

for understanding the mechanisms supporting perceptual development and how they relate to top-down perception in infancy.
© 2024 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by/4.0/).

Introduction

Recent advances in psychology and neuroscience reveal the essential roles of top-down processes in perception (Gandolfo & Downing, 2019; Kok et al., 2017; Oliva & Torralba, 2007; Squire et al., 2013; Summerfield et al., 2006). Top-down processes allow higher-level representations to influence perceptual processing. These influences manifest in the form of transient perceptual changes caused by signals originating from the systems beyond perception (de Lange et al., 2018; Lamme, 2018; Summerfield & de Lange, 2014). The top-down processes in perception have been thought of as an outcome of neural development and emerge only after the neural system matures because the underlying feedback neural connections undergo a prolonged development (e.g., Cao et al., 2017; Dubois et al., 2014). However, this perspective has been challenged by recent imaging and behavioral evidence, revealing the early emergence of the top-down processes in infancy (e.g., Dumont et al., 2022; Emberson et al., 2015; Flaten et al., 2022; Kouider et al., 2015; Xiao & Emberson, 2023). Despite these recent findings of infants' top-down perceptual capacities, the specificity of this cognitive process remains largely unknown. Previous studies have shown that various forms of perceptual specificities, such as advantages in recognizing own-race faces and categorizing other-race faces, arise when infants have relatively homogeneous exposure to face categories and are thought to reflect broader underlying specialization of perceptual systems (e.g., Kelly et al., 2007; Quinn et al., 2016). These specialized perceptual abilities signify the adaptive nature of the perceptual system to the specific characteristics in the sensory environment (for reviews, see Maurer & Werker, 2014; Quinn et al., 2019), and the early-emerged perceptual specificities are often seen as a precursor of perceptual and cognitive abilities. As a newly discovered perceptual capacity in infants, it is entirely unknow how top-down processes adapt to environmental factors such as sensory statistics. To this end, the current study examined infants' engagement in top-down perception with own- and other-race faces.

Perceptual systems in general and the visual system in particular are well-known for their hierarchical neural structure (Gilbert & Li, 2013; Friston, 2005; Lamme et al., 1998; Rao & Ballard, 1999), which progressively processes sensory signals from low- to high-level properties. The upward flow of information processing along the neural hierarchy is referred to as bottom-up processes, which are enabled by feedforward neural connections. In addition to the feedforward connections, a large amount of feedback neural connections in perceptual systems also exist. These connections transmit neural signals from higher-level regions (e.g., the amygdala and prefrontal cortices) or from outside the specific perceptual system (e.g., auditory to visual systems) to modulate perceptual processing as well as from higher- or lower-level regions within perceptual systems. The changes in perceptual processing that arise from information traveling through feedback connections are considered to be top-down perceptual modulations (Emberson, 2017, 2019).

Although the cognitive significance of top-down perception has been revealed in adults (e.g., de Lange et al., 2018; Lamme, 2018), the role of top-down perceptual processing in early development, if any, has yet to be determined (Dehaene-Lambertz & Spelke, 2015). It has been assumed that top-down processes are not available to modulate perception in infancy (Aslin & Smith, 1988; Dehaene-Lambertz & Spelke, 2015) given that they require feedback neural connections, particularly long-range ones, which undergo substantial postnatal change (e.g., Cao et al., 2017; Dubois et al., 2014). The period of development for these top-down connections to be sufficient to support the modulation of perception varies by account from later in infancy (Amso & Scerif, 2015) to childhood (Christiansen & Chater, 2016). There is substantial development of long-range neural connections through adolescence and early adulthood (Sousa et al., 2018; see Bigler, 2021, for a review). However, this view was recently challenged by a series of behavioral and neuroimaging findings that demonstrated the presence of top-down perceptual modulation early in infancy. For example, 6- to 8-month-olds' motion perception can be flexibly modulated by learned auditory cues. When infants faced a directional ambiguous motion display, they perceived leftward motion when they heard a melody associated with left motion. Infants perceived the opposite directional motion when they were exposed to a rightward motion-related cue (Xiao & Emberson, 2023). These convergent findings indicated an early emergence of top-down perception in the first year of life. However, to our knowledge, limited research has addressed how this perceptual ability interacts with infants' experiences that occur outside the lab (“real-world” experiences). Is top-down perceptual processing affected by infants' sensory experiences like other perceptual abilities?

One domain of perceptual development where the role of experience has been extensively investigated is face perception. The experiential influence manifests in the specialization of face perception abilities for infants' commonly seen faces such as own-race faces. For example, after 9 months of age, infants tended to show advanced recognition for own-race faces compared with other-race faces (e.g., Kelly et al., 2007). Experiences with own-race faces also shaped infants' categorization of faces, where infants tended to categorize faces of various other races into a single visual category (Quinn et al., 2016). These specialized face recognition and categorization abilities coincided with the developmental changes in infants' face looking patterns, where infants gradually exhibited distinctive looking patterns for own- and other-race faces (e.g., Ellis et al., 2017; Liu et al., 2011; Wheeler et al., 2011).

Although there has been substantial work investigating infants' perception of own- and other-race faces as well as evidence of top-down perception in infancy, no study to date has investigated infants' top-down perception across own and other races. This lack of research is in part because studying top-down perception is methodologically different from studying perception at a general level. In particular, studies examining infants' top-down perception typically involve cueing or contextual structures that allow infants to anticipate particular stimuli. Top-down perception is evident in contextual effects on perception. Relatedly, to determine whether contextual cues affect the perceptual process, it is crucial to use measures that more closely index perception. Thus, paradigms that allow examinations of top-down perception are methodologically distinct from those that are used to gauge infants' perception (e.g., habituation and visual paired comparison paradigm; see Xiao & Emberson, 2019, for a discussion of some of these distinctions in the context of face perception).

Prior work has suggested differential top-down perception of own-race versus other-race faces. Vogel and colleagues (2012) examined whether an emotional sound preceding a face would modulate event-related potential components (N290 and P400) associated with infants' face perception. Related to the current study, they found that the emotional sound modulated the N290 and P400 only with own-race faces but not with other-race faces at 9 months of age. This suggests that top-down perception might be available only for own-race faces. However, the behavioral significance of this face-race-specific neural signature was scarcely evaluated. Xiao and Emberson (2019) examined the impact of emotional sounds on face perception, as measured behaviorally, and found evidence of top-down modulation after 6 months of age. However, they investigated only own-race faces, leaving the specificity of infants' top-down perceptual ability unclear.

Based on the large body of studies on infants' specialized face perception and past work examining the top-down perception of own-race faces, the current study is the first to compare the top-down perception of own- and other-race faces to examine the specificity of infants' top-down perceptual ability. We specifically examined three possible experiential outcomes, namely that (1) infants may engage in top-down perception only with own-race faces but not with other-race faces; (2) infants might rely more on top-down processes to perceive other-race faces as a compensatory mechanism to the inferior perception of other-race faces; and (3) top-down perception could be considered a general cognitive construct, unaffected by sensory experiences, and therefore it could be equally applicable to faces of both one's own race and other races.

Our first hypothesis proposed an advanced top-down perception with own-race faces in comparison with that with other-race faces. Specifically, we proposed that the substantial experience that infants gain with own-race faces could provide opportunities to engage in top-down perception, strengthening top-down processing for own-race faces. There are many ways that this experience-based increase in top-down processing could occur. The increased representation or perceptual processes in relation to own-race faces can result in better top-down processing. Another not mutually exclusive possibility is that infants have many more opportunities to engage in top-down processing of own-race faces (e.g., infants could use someone's voice to augment their visual perception of the related face), and these experiences may optimize the feedback neural networks to support the top-down perception of own-race faces, resulting in poorer performance for other-race faces. This hypothesis is similar to a proposal by Markant and Scott (2018), which emphasized that the attentional systems developed alongside face specialization in infancy result in specialized attention for familiar or own-race faces.

Alternatively, prior research suggests that top-down processing may be strongly involved in the perception of difficult-to-perceive or unfamiliar stimuli, and thus top-down processing would be more effective in other-race faces than in own-race faces. Converging evidence of top-down perception from infants to adults has shown that top-down processes play a compensatory role in perception. Top-down processes were salient when perceptual signals were weak or when tasks were challenging (e.g., Bar, 2004; Dowdle et al., 2021; Hupé et al., 1998; Xiao & Emberson, 2023). For example, in a recent study on infants' motion perception, top-down modulation was significant only when motion signals were ambiguous. When very strong motion signals were presented, no evidence of top-down processing was found (Xiao & Emberson, 2023). In infancy, the lack of experience leads to difficulties in perceiving unfamiliar stimuli. Infants might rely on top-down processes to overcome these perceptual challenges. In line with this prediction, several recent studies have shown that when infants learned other-race faces accompanied by auditory or expressive information, which could act as sources of top-down modulation, their face recognition performance improved significantly (Minar & Lewkowicz, 2018; Quinn et al., 2020). However, whether the top-down process was enhanced during the perception of unfamiliar face categories remains untested because there was no comparison with familiar face categories.

A third possibility is that there is no difference between types of stimuli. This pattern of results may arise if top-down perception is supported principally through age-related changes in the maturation of long-range connectivity that is independent of experience. If this is the case, changes in top-down perception should be equal across different types of stimuli.

To probe the specificity of infants' top-down perception, the current study focused on their perception of rapidly presented faces. Infant participants saw two streams of rapidly and serially presented visual masks on each side of the screen. In one stream, a face image embedded in this stream of visual masks would appear briefly ($\leq$200 ms). Infants' perception of the face was indexed by their looking preference for the face side over the exclusively visual mask side. This paradigm has been used to assess infants' face perception from 5 to 15 months of age (Gelskov & Kouider, 2010). To reveal the top-down modulation, we presented two types of audio sounds along with the visual display. One sound predicts faces, but the other does not (i.e., it predicts another visual category: flowers). The top-down effect would be manifested by the difference in face perception between the predictive and irrelevant sound conditions. Specifically, we expected a larger preferential looking to the face side in the predictive sound condition than that in the irrelevant sound condition.

There are several key advantages to the use of this paradigm for measuring infant face perception compared with traditional methods (e.g., habituation, visual paired comparison). First, this task measured infant perception in real time when perception occurred and did not rely on memory. Moreover, within this paradigm, perception could be quickly assessed (i.e., within seconds rather than minutes). This rapid assessment of perception allowed us to use a within-participant design to examine how infants' face perception could be flexibly modulated by top-down signals. In other words, we could compare the perception of faces across multiple perceptual conditions when infants were hearing predictive sounds versus when they were not. Furthermore, given the fact that this paradigm can manipulate the amount of sensory input (i.e., the face presentation duration), we could examine how much sensory input was required for infants to engage in top-down perceptual modulation. Previous findings suggest that top-down modulation occurred when bottom-up sensory input was weak or ambiguous (e.g., Bar, 2004; Dowdle et al., 2021; Hupé et al., 1998). To this end, we selected several face presentation durations, which allowed us to examine the circumstances in which the top-down effect emerged in the current paradigm and how it related to the sensory availability of the face image.

In two experiments, we investigated top-down perceptual modulation on perceiving own-race faces (Experiment 1) and other-race faces (Experiment 2). It is worth noting that infant participants were 10 to 13 months old and therefore were expected to show race-based narrowing in face perception given relative homogeneity in community diversity exposure (i.e., perceptual narrowing from 6 to 9 months of age; Maurer & Werker, 2014) and thus have been shown to exhibit the other-race effect (i.e., reduced processing of other-race faces compared with own-race faces).

Experiment 1

Method

Participants

A total of 29 White infants (12 girls) participated in the current experiment. The average age was 355 days (*SD* = 28.06), ranging from 306 to 397. The number of participants in the current study was determined based on our power analysis (see online supplementary material). All participants were White from monoracial White families recruited from a city with a high percentage (69%) of White people. Thereby, their face experience was dominated by White people, and we assumed that these infants had a high degree of perceptual capacity for the White faces they would perceive in the current experiment. All participants were full-term infants with normal vision and hearing. An additional 12 infants participated in the study but were not included in the final sample due to fussiness during the experiment ($n = 6$) or their inability to provide data for at least one of the face sound versus flower sound comparisons (100, 150, or 200 ms; $n = 6$).

Materials and procedure

The institutional review board of Princeton University approved all study procedures. The study comprised multiple experimental blocks. Each experimental block included three types of sequentially presented phases: association learning phase, face perception test phase, and learning validation phase. Experimental blocks were repeated in this sequence until infants became fussy or finished 6 blocks (Fig. 1).

In the association learning phase, infants learned two audio–visual associations: One sound was paired with face images, and another sound was paired with flower images. In this phase, infants saw a face presented on the left or right side of a computer screen for 250 ms, which was followed by a 750-ms blank screen. Beginning during the face presentation, infants heard a corresponding auditory stimulus (*face sound*) playing for 1 s. Similarly, another auditory stimulus (*flower sound*) was played when infants saw flowers in the flower association learning trials. The location of face and flower images was counterbalanced across trials so that there was no spatial contingency between either the sounds or the type of visual stimulus. The face sound and flower sound were randomly selected from two instrumental chords (tuba F & guitar C#) for each participant. To induce the learning of the audio–visual associations, we designed 24 learning trials (12 face learning trials and 12 flower learning trials) in the first block. At the beginning of the subsequent blocks, we placed 16 learning trials (8 face learning trials and 8 flower learning trials) to consolidate the learning.

After the association learning phase, the top-down influence on face perception was examined with the face perception test phase. Following the procedure used by Gelskov and Kouider (2010), infants saw two streams (*mask stream* and *face stream*) of images sequentially presented on both the left and right sides of the screen. The mask stream included 14 different visual masks. The face stream was identical to the mask stream except that the 2nd, 5th, 8th, and 11th images were an image of the same face. The face stream was presented on the left side for half of the test trials and on the right side for the other half. All images were displayed briefly on the screen. For the face stream, the face was presented for one of three durations (100, 150, or 200 ms) within a given test trial. The visual mask that immediately followed the face images (i.e., the 3rd, 6th, 9th, & 12th images in the stream) was presented for 33 ms. The images on the mask stream were presented simultaneously with those in the face stream except that no faces were presented. Thus, participants saw the image in both streams changing simultaneously.

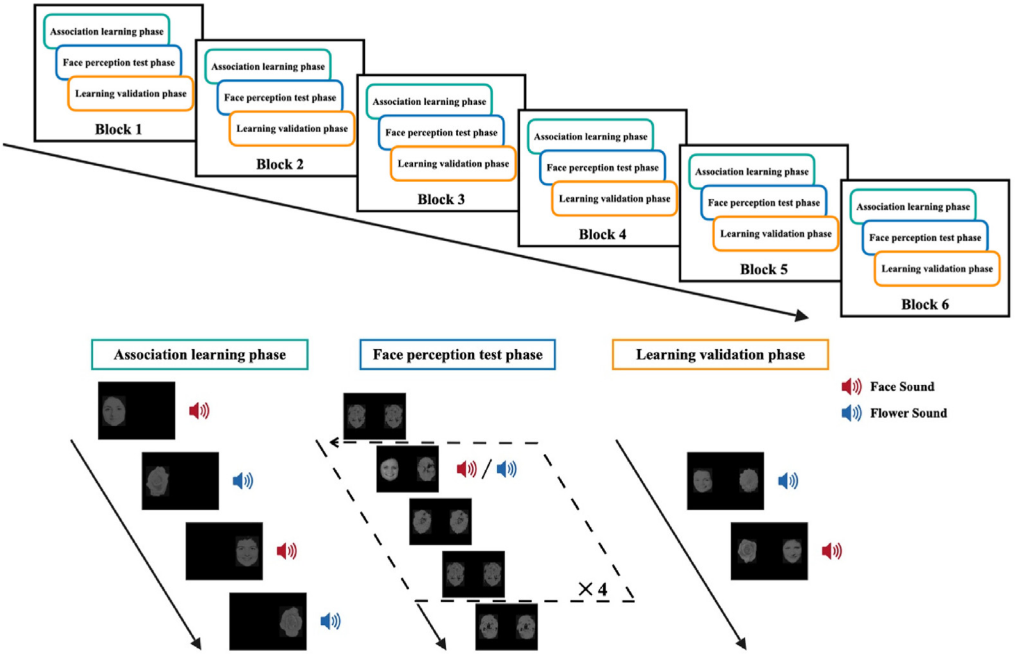

Fig. 1. Schematic presentation of the experimental procedure (top panel) and experimental procedures for the association learning phase, face perception test phase, and learning validation phase (bottom panel) in Experiment 1.

The visual masks were composed of scrambled pixels of the face images and had the same contours as those face images. Thus, the masks and faces matched the overall brightness, contrast, and other low-level perceptual features as such. The inclusion of these masks reduced the possibility of detecting faces by the contour or complexity of visual elements. All images were 583 × 750 pixels presented at the left side (horizontal: 320 pixels, vertical: 512 pixels) and right side (horizontal: 960 pixels, vertical: 512 pixels) of a 1280 × 1024-pixel screen (17 inches). With a viewing distance of 60 cm, the visual angle of each image was 14.83° horizontally and 19.40° vertically.

A sound would play continuously as infants watched the two streams of images. The sound was either the face sound (predictive sound) or the flower sound (irrelevant sound) that infants learned in the association learning phase (Fig. 2, left panel). As in Gelskov and Kouider (2010), the same face image would appear again at the end of each face perception test trial with a rewarding sound, which was randomly selected from six amusing sound clips (e.g., sounds of bubbles) for 3 s to reinforce infants' looking at the streams.

Notably, face perception test trials varied the duration of face presentation to manipulate the availability of a bottom-up perceptual signal. The face was presented for one of three durations (100, 150, or 200 ms) in each test trial, with the duration constant within trials. The reason for choosing three duration levels was to vary the perceptual difficulty of the task. These three duration levels were chosen because infants' face perception threshold falls between 100 and 200 ms, as measured with this paradigm (Gelskov & Kouider, 2010). Specifically, the 100-ms duration is likely under infants' face perception threshold, whereas a face presented for 200 ms should be at or above their perception threshold. Should infants engage in top-down processes, the three duration levels would allow us to examine under what circumstance the top-down effect emerges with the current paradigm and how it relates to sensory availability of the face image. To this end, there were 6 types of test trials within each block: a total of 2 sounds (face sound and flower sound) × 3 presentation durations (100, 150, and 200 ms) = 6 trials presented in random order.

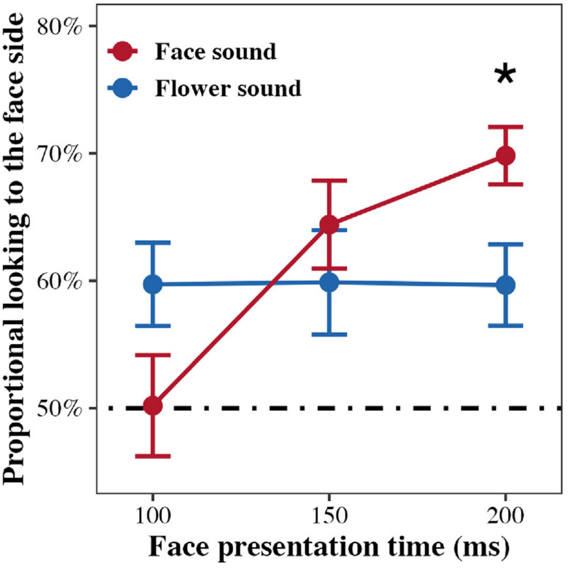

Fig. 2. Mean proportional looking to the face side during the rapid face presentation on the test trials in Experiment 1. Asterisk indicates significant difference in the proportional looking between face sound trials and flower sound trials when faces were presented for 200 ms. Error bars represent 1 standard error.

After the face perception test phase, we examined infants' learning of the associations with 2 learning validation trials in the learning validation phase. In each trial, a randomly selected face image and a randomly selected flower image were presented side-by-side for 2.75 s. The face sound was played in 1 of the trials, and the flower sound was played in the other trial. The locations (left vs. right) of the face and flower images were counterbalanced across the 2 trials. These learning validation trials were designed to use infants' looking preferences to examine the strength of the learned associations. Infants were expected to look longer at the image associated with the corresponding sound should they learn the associations.

All images were grayscale images presented against a black background. We downloaded them from the internet. We prepared 22 images of smiling White female faces and 22 images of flowers. To reduce the impact of individual face images on infants' learning and perception, each face image appeared in only 1 trial and would never repeat across trials within a block. This design of the test trial allowed us to determine whether infants had learned the association between the sounds and a visual category (i.e., face or flower) rather than individual exemplars from the category.

The experiment was controlled by Psychtoolbox (3.0.13) on MATLAB. Infants' eye movement was recorded throughout the experiment by the EyeLink 1000 (16-mm lens) system with a 500-Hz sampling rate. A five-point (four corners and screen center) calibration was performed before the experiment. The calibration quality was assessed via the same five-point validation procedure provided by EyeLink.

Results and discussion

Top-down modulation on face perception

On average, participants finished 3 learning blocks (2–5 blocks). Infants' face perception was measured based on their eye movement data recorded in the face perception test phase. Specifically, face perception was operationalized by the proportional looking time at the side where the face was presented. The areas of interest for each screen side were the left and right halves of the screen. A proportional looking rate above 50% indicates that infants could perceive the rapidly presented face.

The larger the proportional looking, the better infants perceived the rapidly presented face (Gelsov & Kouider, 2010). A top-down perceptual effect would manifest as larger proportional looking to the face side in the face sound condition than that in the flower sound condition.

Participants finished an average of 18 test trials (10–30 trials). To evaluate the presentation dependency of the top-down effect, we performed a 2 (Predictive Sound: face vs. flower) × 3 (Face Presentation Duration: 100, 150, or 200 ms, continuous) repeated-measures analysis of variance (ANOVA) on the proportional looking at the face side. This analysis focused on the participants who generated valid eye movement data for all six conditions (3 presentation times × 2 sounds). Of the 29 participants in Experiment 1, 24 met this criterion, and their data were analyzed. We found a significant interaction, $F(1, 23) = 12.44$, $p = .002$, $\eta_p^2 = .15$, indicating that the top-down processing depended on the availability of bottom-up information. Moreover, the ANOVA also revealed a significant main effect of face presentation durations, $F(1, 23) = 15.86$, $p < .001$, $\eta_p^2 = .15$, indicating that the face perception improved with the duration of face presentation, which was consistent with previous studies (e.g., Gelsov & Koudier, 2010). No main effect of predictive sound was found, $F(1, 18) = 0.45$, $p = .509$, $\eta_p^2 = .01$.

To further explore the significant interaction, we did a post hoc analysis to compare the looking proportion at the face side between the face sound and flower sound conditions with a paired-sample *t* test at each presentation time condition (100, 150, and 200 ms), respectively. For faces presented at 100 ms, we did not find a significant difference in infants' face preferential looking to the face side between the face sound condition ($M = 50.19\%$, $SD = 19.45$) and the flower sound condition ($M = 59.72\%$, $SD = 15.97$), $t(23) = -2.01$, $p = .056$, Cohen's $d = 0.41$. Infants also showed similar preferential looking with 150 ms presentation duration in the face sound condition ($M = 64.40\%$, $SD = 16.90$) and flower sound condition ($M = 59.88\%$, $SD = 20.10$), $t(23) = 0.833$, $p = .413$, Cohen's $d = 0.17$. In contrast, when faces were presented at 200 ms, infants exhibited significantly larger preferential looking at the face side in the face sound condition ($M = 69.81\%$, $SD = 11.05$) than in the flower sound condition ($M = 59.66\%$, $SD = 15.62$), $t(23) = 2.47$, $p = .021$, Cohen's $d = 0.50$. These results indicated that the top-down influence on infants' face perception emerged only at the greatest level of sensory face information in this task (200 ms).

It should be noted that the ANOVA and post hoc analyses included only the participants who contributed valid looking data to all six conditions. In addition to these children, 5 participants contributed valid looking data to part of the conditions. Although we could not include these participants in the ANOVA, they were eligible for the *t* tests comparing the looking proportion between the face sound and flower sound conditions at each presentation level as well as the following analysis that examined participants' associative learning. By including these additional participants, the statistical results were expected to be more robust. The results are reported in the supplementary material.

Association learning performance

We examined whether infants learned the association between a given sound and its corresponding visual category (e.g., the association between the face sound and faces) by analyzing their looking preferences in the learning validation phase. The proportional looking time to the stimulus matching the sound presented in each trial (i.e., looking proportion to the face side when the face sound was played and looking proportion to the flower side when the flower sound was played) was calculated. Participants finished 3 learning validation trials (1–5 trials). We found that infants looked mostly at the stimuli related to the sound ($M = 58.88\%$, $SD = 11.31$), which was significantly different from the chance level (50%; $n = 28$), one-sample $t(27) = 4.15$, $p < .001$, Cohen's $d = 0.79$. This result indicated that infants learned the audio–visual associations.

Experiment 1 examined how learned predictive cues (the face sound) affected infants' perception of rapidly presented own-race faces. This experiment was built from the previous work that established this paradigm (Gelsov & Kouider, 2010). Without considering the effect of the different audio cues, the current study replicated previous findings regarding infants' face perception: With the increased face presentation duration, infants exhibited increased looking time toward the side where faces appeared. This finding indicated that the amount of bottom-up sensory input affected infants' face perception.

This experiment also examined how the presence of predictive top-down cues affected infants' face perception through the presence of sounds that either predicted or were associated with faces or another visual category, namely flowers. Expanding on previous studies that used emotional vocal cues (Xiao & Emberson, 2019), the current experiment placed infant participants in a scenario where they could use an audio cue to interact with their visual perception of faces. We found the modulation effect of audio on face perception; infants exhibited significantly better perception of faces (i.e., looking time to the face) when they heard the sounds that were predictive of faces as opposed to when they heard the sound that was irrelevant to faces. Because the only difference between the two conditions was the predictiveness of the sounds, the observed perceptual effect must be attributed to the audio cues. Thus, this effect demonstrated a top-down modulation of face perception in infants.

Lastly, we found that the top-down effect was dependent on the face presentation duration. The top-down effect emerged when substantial bottom-up information (200 ms presentation duration) was available. The effectiveness of the top-down modulation, at least at the behavioral level, required a certain amount of sensory input.

Experiment 2

Having determined that infants used top-down signals to boost their face perception, it remained unclear whether this capacity is specific to own-race faces. To address this research question, Experiment 2 investigated infants' top-down effect on the perception of other-race faces.

Method

Participants

A total of 27 White infants (10 girls) participated in the current experiment. The average age was 339.52 days (*SD* = 27.14), ranging from 294 to 380. All participants were White from monoracial White families recruited from a city with a high percentage (69%) of White people. All participants were full-term infants with normal vision and hearing. An additional 16 infants participated in Experiment 2 but were excluded from the final sample because of fussiness during the experiment (*n* = 7) or their inability to provide data for at least one of the face sound versus flower sound comparisons (100, 150, or 200 ms; *n* = 9). We also surveyed infant participants' daily experience with Asian people and found that 2 infants had constant interactions with Asian people (i.e., daily interactions with Asians for more than 1 h). Thus, we removed these 2 infants from the following analyses. The remainder either had sporadic interactions with Asians (*n* = 4) or lacked any exposure to Asians (*n* = 23). None of the infants in this experiment participated in Experiment 1. Informed consent was obtained for the current experiment with human participants, consistent with the institutional review board procedures noted above.

Materials and procedure

All faces used in Experiment 2 were smiling Asian faces, which were processed in low-level properties such as brightness, contrast, and size to best align with those of White faces used in Experiment 1. Except for changing the face stimuli, the experimental procedure of Experiment 2 was identical to that of Experiment 1.

Results and discussion

Face detection

On average, participants finished an average of 4 learning blocks (3–6 blocks), 20 test trials (16–34 trials), and 5 learning validation trials (3–8 trials). We performed the same 2 (Predictive Sounds: face vs. flower) × 3 (Presentation Duration: 100, 150, or 200 ms) repeated-measures ANOVA on the proportion of looking to the face side in the face perception test phase. As shown in Fig. 3, the results only revealed a significant main effect of presentation durations, $F(1, 26) = 4.22$, $p = .050$, $\eta^2 = .06$. However, no main effect of the predictive sound was found, $F(1, 26) = 0.10$, $p = .756$, $\eta^2 = .001$, and no significant

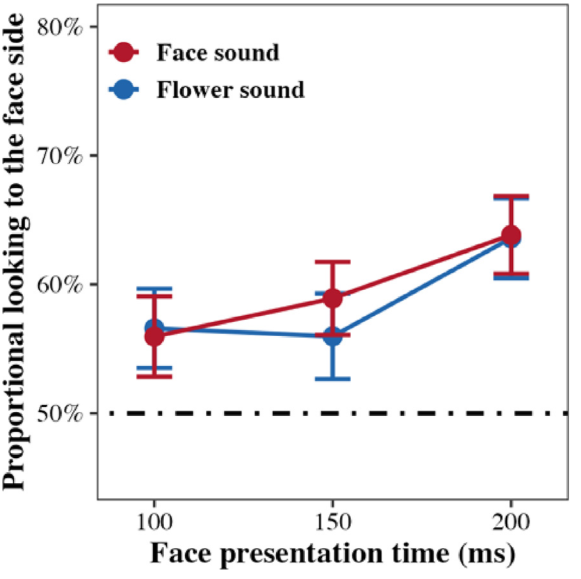

Fig. 3. Mean proportional looking to the face side during the rapid face presentation on the test trials in Experiment 2. Error bars represent 1 standard error.

interaction was found, $F(1, 26) = 0.03$, $p = .872$, $\eta^2 < .001$. In sum, these findings suggested that although infants' face detection performance improved with the increase of face information (as found in Experiment 1 and Gelsov & Kouider, 2010), face perception did not benefit from the learned predictive cues, thereby proving no top-down modulation. The failure to find an interaction between sound type and presentation durations suggested that the top-down modulation on face perception is specific to own-race faces.

The contrast in the top-down effect between own- and other-race faces may help to reveal the mechanism of the top-down effect regarding whether the predictive face sound boosted face perception or the irrelevant flower sound inhibited face perception. Because no top-down modulation was found in the perception of other-race faces, the observed other-race perception may serve as a baseline, representing infants' face perception in the absence of the top-down influence. Thus, by comparing perception performance between the own- and other-race conditions, we can infer whether top-down signals boost the perception of related information in the predictive face sound condition and whether the irrelevant information perception is inhibited by top-down modulation in the flower sound condition. The comparisons showed a descriptive advantage of top-down signals in processing relevant information, independent-sample $t(49) = 1.56$, $p = .125$, Cohen's $d = 0.44$, and a descriptive decrease in face perception of irrelevant top-down signals, independent-sample $t(49) = -0.88$, $p = .383$, Cohen's $d = 0.25$. Although these effects failed to reach statistical significance, the descriptive differences imply that top-down signals boosted the processing of relevant perceptual information while inhibiting the processing of irrelevant information. However, the comparisons were exploratory and not planned by the current design. The current sample size also may be too small to reveal the true effect or this effect could be spurious. Future studies may consider a more specific and intentional design (e.g., having positive, negative, and neutral top-down signals with a larger sample size) to either confirm or disconfirm these observations and examine the specific roles of top-down signals.

Equivalent associative learning across own- and other-race faces

Given that the top-down cues used in the current paradigm were learned within the task, it is important to ensure that the differences in the top-down effect between the two experiments were not arising from differences in infants' ability to form an association between a sound and faces across own- and other-race face categories. Specifically, infants might have a poorer ability to learn the association between a sound and other-race faces than between a sound and own-race faces. Recent studies on infants' statistical learning abilities showed that infants are more apt to learn associations and regularities with stimuli from familiar categories (e.g., upright faces, own-race faces) as opposed to unfamiliar ones (e.g., inverted faces, other-race faces) (Bulf et al., 2015; Santolin & Saffran, 2018; Xiao et al., 2018). If there were differences in the associative learning necessary for top-down processes, these differences in top-down effects may arise from differences in learning rather than the use of learned information to modulate perception. Therefore, our experimental design included the learning validation phase to evaluate infants' associative learning performance.

To ascertain whether the variations in top-down modulation resulted from differences in the associative learning of audio cues with faces, we analyzed infants' looking preferences in the learning validation phase across experiments. Contrary to this alternative account, infants showed significant looking to the stimulus indicated by the sound (M = 55.44%, SD = 12.15), one-sample t test, $t(26) = 2.33$, p =.028, Cohen's d = 0.45, in Experiment 2, as in Experiment 1. Notably, as shown in Fig. 4, infants' looking preference in Experiment 2 was not different from that in Experiment 1, independent-sample $t(53) = 1.09$, p =.282, Cohen's d = 0.29, providing evidence that infants' learned association between other-race faces and a sound was not different from their learned audio–visual association with own-race faces. Thus, this finding provided evidence that, in this paradigm we were able to circumvent the differences in statistical and associative learning previously observed between own- and other-race faces and, instead, were able to uncover additional differences in the use of this learned information to modulate face perception based on top-down cues.

General discussion

Challenging the idea that the infant perceptual system is largely bottom-up (e.g., Aslin & Smith, 1988; Cao et al., 2017), recent studies have shown evidence of top-down modulation of perception in infancy using both behavioral and neural measures (e.g., Emberson et al., 2015; Kouider et al., 2015; Vogel et al., 2012; Xiao & Emberson, 2019, 2023). Having established the presence of top-down mechanisms in infancy, it is now important to understand whether and how these top-down

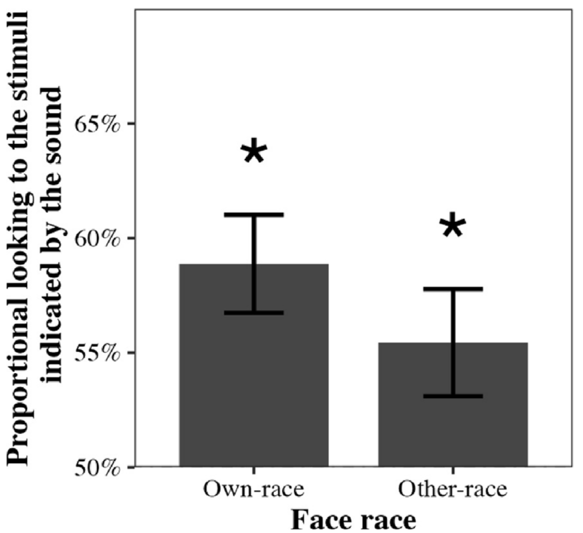

Fig. 4. Mean proportional looking to the stimuli indicated by the sound during learning validation trials for own-race faces (Experiment 1) and other-race faces (Experiment 2). Asterisks indicate significant differences between looking proportion and chance level (50%) with one-sample t tests. Error bars represent 1 standard error.

mechanisms contribute to perceptual development. With two experiments, the current study investigated the specificity of infants' ability to engage in top-down face perception. We found that infants aged 10 to 13 months can modulate the perception of rapidly presented own-race faces by using recently learned predictive audio cues. Face perception was augmented by predictive cues (i.e., ones that they learned were associated with faces) as opposed to irrelevant cues (i.e., ones that they learned were associated with flowers). To our knowledge, these findings provide the first evidence that infants' capacity to engage in top-down modulation of face perception may become specialized for own-race faces after 10 months of age. The current findings suggest that the specialization of perceptual systems in infancy could be influenced by the specialization of top-down processes from higher-level systems.

Considering why top-down processes differed between the own- and other-race conditions, we can rule out one of the most likely reasons for these differences. First, previous work has demonstrated that infants have an advantage in learning associations and regularities with categories of familiar stimuli (e.g., upright faces, own-race faces) as opposed to unfamiliar ones (e.g., inverted faces, other-race faces) (Bulf et al., 2015; Santolin & Saffran, 2018; Xiao et al., 2018). However, in the current study, to validate infants' learning, we found that they learned the associations of own- and other-race faces significantly and equally well. The learning conditions in the current study were purposefully easy, so we do not interpret this lack of difference between face types as evidence to counter these prior studies but, rather, to assert that this difference in learning is not the origin of the difference in top-down modulation. Moreover, we can rule out that there are base-level differences in infants' perceptual abilities in this task (independent of top-down modulation). Specifically, we found that infants' perceptual response to the availability of bottom-up signals (i.e., duration of face presentation) was similar between own- and other-race faces (i.e., the main effects of presentation durations). Thus, it is not the case that the perceptual conditions were so poor for the other-race faces that there was no perceptual signal to be modulated by top-down signals. Together, with respect to the current paradigm, these results suggest that the difference between own- and other-race face perception lies in the capacity to engage top-down processes.

Future work is needed to determine how these top-down mechanisms become specialized to familiar stimuli. Here, we propose several possibilities. One possible route is that infants' experiences with their own-race faces shape the development of feedback neural connections, which are the neural foundation for top-down modulation. Experience with own-race faces provides infants with opportunities to learn statistical associations between own-race faces and other signals such as the relation between voices and faces (e.g., Kubicek et al., 2014). It is already established that infants can use previously learned associations as sources of top-down signals to assist perceptual processing, provided that they are in contexts to use them (see Xiao & Emberson, 2019). Thus, the more own-race face experience infants have, the more frequently they can engage in top-down modulation on own-race face processing. As a result, the feedback neural network that supports top-down modulation on own-race face processing may be gradually strengthened, leading to more effective transmission of top-down signals and modulations on perceiving own-race faces. In this route, top-down mechanisms exhibit differential development for own- and other-race faces based on top-down modulation across months of experience.

Another non-mutually-exclusive reason for the observed race effect points toward the differences in face perception as opposed to developmental differences in top-down processing per se. Previous work has argued that an infant's experiences with own-race faces lead to specialization in face perception systems, with one of the consequences being the difficulty in perceiving unfamiliar face categories (see Maurer & Werker, 2014, for a review). The differences in the representational fidelity of these different faces may result in varying effectiveness when engaging in top-down processes to modify perception. Thus, the same strength of top-down signals could be applied to both cases, but perhaps these top-down signals are better able to modify the perceptual representations with greater fidelity (i.e., own-race faces). In this view, top-down mechanisms are more effective for faces with more specialized or finer representations. This is an important avenue for future investigation: Does the representational fidelity of a given stimulus affect how strongly the top-down signal can modify its perception? However, this idea, at least on the surface, runs counter to work suggesting that top-down signals are most effective when bottom-up perceptual signals are weak (e.g., Bar, 2004; Hupé et al., 1998). Thus, if the representational fidelity of own- versus other-race faces can be equated to bottom-up perceptual signals, we may expect greater benefit for top-down signals for other-race faces.

A potential way to distinguish the two hypotheses (Fig. 5) is to focus on whether the perception of own- and other-race faces can be modulated by a predictive cue, which is not specific to any face race. For example, in contrast to the current study where the learned cue was associated with either own-race faces (Experiment 1) or other-race faces (Experiment 2), participants were trained to associate a cue with faces irrespective of their race, based on their experience with both own- and other-race faces. Then, we can examine how this learned cue modulates the perception of own- and other-race face perception. In this setting, the strength of the top-down signal should be equivalent for both own- and other-race faces. Under the first condition, where this difference arises from the top-down signal, we should expect comparable top-down modulation between the own- and other-race conditions. However, if discrepancies in the representations of own- versus other-race faces are driving this difference, we should continue to find an advantage for own- versus other-race faces, which is similar to the current finding. This could be an important avenue for future work to investigate the relation between top-down signals themselves, the representational changes in face perception, and their interaction.

These findings dovetail with recent theoretical advances in early perceptual development (e.g., Markant et al., 2016; Markant & Scott, 2018). Markant and Scott (2018) proposed that there is a bidirectional relation between selective attention (one possible top-down mechanism) and face perception. The interactions between the two systems result in the specialization of both attentional and perceptual systems to more experienced own-race faces. More specifically, increased attentional weighting of features on own-race faces (e.g., greater sustained attention, faster orienting) converges with increased specialization of posterior perceptual cortices for these facial features. Broadly, the current findings are compatible with Markant and Scott's (2018) perspectives. One major difference is that whereas Markant and Scott focused on selective attention as a source of top-down modulation, we explicitly considered a potentially large number of sources of top-down processes for perceptual systems beyond selective attention. For example, Xiao and Emberson (2019) argued that emotional information, which likely arises from feedback from the amygdala to the fusiform face area, modulated face perception by 10 months of age. In the current study, the top-down signals likely originated from associative learning systems (e.g., the prefrontal cortex; Werchan et al., 2016). These recent advances in infant perception highlighted the existence of a wide range of top-down mechanisms across infants' cognitive systems. It is an open question whether these are importantly different

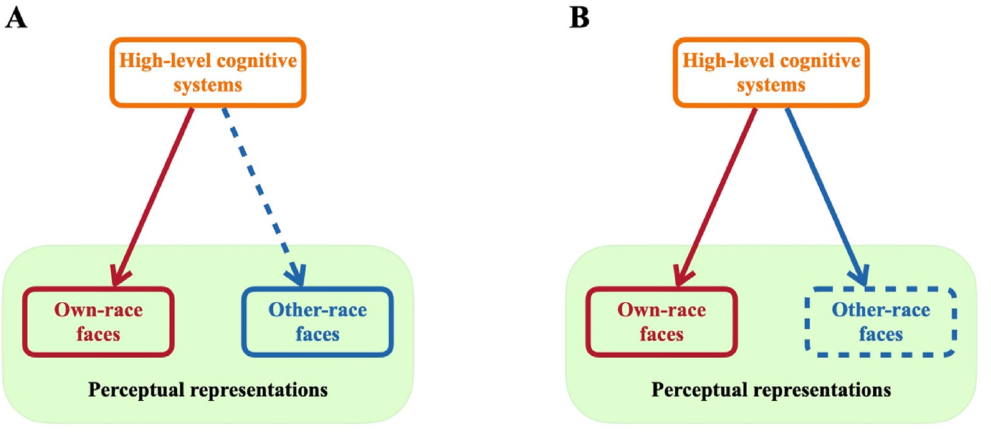

Fig. 5. Illustrations for the hypotheses for the own-race-specific top-down perceptual modulation. Panel A demonstrates that the feedback neural connections (solid vs. dashed straight lines) from high-level cognitive systems to perceptual systems become specialized toward own-race faces. Panel B demonstrates that the fidelity of representations (solid vs. dashed line borders) for own- and other-race faces determines the effectiveness of top-down modulation.

mechanisms or similar. Do they all show sensitivity to experience, as demonstrated here, or not? If so, what differences are relevant? Indeed, more recent work has suggested that it is not under all circumstances that selective attention shows biases in relation to the race of faces (e.g., Hunter & Markant, 2021; Prunty et al., 2020).

This specialized top-down process enhances our understanding of early development by highlighting the importance of experience. Currently, evidence for the development of infants' top-down perception is revealed exclusively through changes across ages (Nakashima et al., 2021; Xiao & Emberson, 2019). For example, Nakashima and colleagues (2021) examined the development of recurrent connections in the visual system using a classical paradigm known as object substitution masking. In this paradigm, top-down perception is demonstrated by the disruption of face perception caused by a persistent mask following face presentation. The authors found a disruption only in 7- and 8-month-old infants but not in younger infants aged 3 to 6 months. Together with the current findings, this evidence strongly suggests that early infancy is a critical period for the development of top-down processes in perception. However, to fully understand the developmental mechanisms of top-down perception, we need to further examine the emergence of specialized top-down capacity throughout infancy. In particular, researchers need to use a paradigm more suitable for young infants. The current paradigm requires the mature oculomotor ability to move eyes quickly between two sides of the screen, which is often challenging for infants younger than 5 months when this ability is underdeveloped. Moreover, by tracking the top-down perception of other-race faces in toddlers and young children, we would know whether the top-down processes in the perception of unfamiliar types of faces undergo a delayed development or are absent.

One limitation of the current study is that we recruited only White infants because the study was conducted in a city with a high percentage (69%) of White people. The observed race effect may represent stimuli differences between the White face and Asian face images rather than differences in modulating the perception of own- and other-race faces. Given the fact that no systematic difference between perceiving White and Asian faces was reported (i.e., similar overall looking percentage to the face side) in the current study, we do not think that visual differences between the two races of faces would lead to the current finding. Nevertheless, it would be invaluable to replicate the current findings with infants from other racial backgrounds.

To summarize, the current study showed that infants' face perception could be modulated by the amount of bottom-up sensory input and the top-down signals engendered by learned and predictive audio cues. The two sources of signals (bottom-up and top-down) work in tandem to optimize infants' perceptual processing when confronted with dynamic environments. Although the two sources of signals contribute to infants' perception of own-race faces, the top-down signals were not found to modulate the perception of other-race faces. The finding of differential top-down effects for own- versus other-race faces suggests that top-down mechanisms specialize relatively early in development (i.e., 10 months of age). Moreover, these top-down mechanisms selectively offer flexible and context-dependent modulation toward commonly experienced stimuli in infants' environments. These findings extend our current understanding of early perceptual development and reveal that it not only is driven by experience-based structural changes within perceptual systems but also is integrated with higher-level cognitive systems such as learning, memory, and attention. Moreover, this article suggests that it is not simply perceptual systems that specialize; systems that can provide top-down signals also exhibit specialization.

CRediT authorship contribution statement

Naiqi G. Xiao: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Data curation, Conceptualization. **Hila Ghersin:** Data curation. **Natasha D. Dombrowski:** Data curation. **Alex M. Boldin:** Software. **Lauren L. Emberson:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Conceptualization.

Data availability

All raw data were uploaded to the Open Science Framework (https://osf.io/m6gxe/?view_only=55d96b8374d84cfc851c226a49375f8b). These data and scripts are accessible to anyone.

Acknowledgments

This research was supported by grants from the National Institutes of Health (R00 4R00HD076166-02) and McDonnell Foundation (220020505) and by Natural Sciences and Engineering Research Council of Canada Discovery Grants (RGPIN-2020-07129). We thank Rachel (Xinyang) Liu, Emma Resendes, Daniel Lee, and Grant Bae for assisting in proofreading the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2024.105889.

References

Amso, D., & Scerif, G. (2015). The attentive brain: Insights from developmental cognitive neuroscience. *Nature Reviews Neuroscience, 16*(10), 606–619. https://doi.org/10.1038/nrn4025.

Aslin, R. N., & Smith, L. B. (1988). Perceptual development. *Annual Review of Psychology, 39*(1), 435–473. https://doi.org/10.1146/annurev.ps.39.020188.002251.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*(8), 617–629. https://doi.org/10.1038/nrn1476.

Bigler, E. D. (2021). Charting brain development in graphs, diagrams, and figures from childhood, adolescence, to early adulthood: Neuroimaging implications for neuropsychology. *Journal of Pediatric Neuropsychology, 7*(1–2), 27–54. https://doi.org/10.1007/s40817-021-00099-6.

Bulf, H., Brenna, V., Valenza, E., Johnson, S. P., & Turati, C. (2015). Many faces, one rule: The role of perceptual expertise in infants' sequential rule learning. *Frontiers in Psychology, 6*, 1595. https://doi.org/10.3389/fpsyg.2015.01595.

Cao, M., Huang, H., & He, Y. (2017). Developmental connectomics from infancy through early childhood. *Trends in Neurosciences, 40*(8), 494–506. https://doi.org/10.1016/j.tins.2017.06.003.

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences, 39*. e62. https://doi.org/10.1017/S0140525X1500031X.

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How do expectations shape perception? *Trends in Cognitive Sciences, 22*(9), 764–779. https://doi.org/10.1016/j.tics.2018.06.002.

Dehaene-Lambertz, G., & Spelke, E. S. (2015). The infancy of the human brain. *Neuron, 88*(1), 93–109. https://doi.org/10.1016/j.neuron.2015.09.026.

Dowdle, L. T., Ghose, G., Ugurbil, K., Yacoub, E., & Vizioli, L. (2021). Clarifying the role of higher-level cortices in resolving perceptual ambiguity using ultra high field fMRI. *NeuroImage, 227*. 117654. https://doi.org/10.1016/j.neuroimage.2020.117654.

Dubois, J., Dehaene-Lambertz, G., Kulikova, S., Poupon, C., Hüppi, P. S., & Hertz-Pannier, L. (2014). The early development of brain white matter: A review of imaging studies in fetuses, newborns and infants. *Neuroscience, 276*, 48–71. https://doi.org/10.1016/j.neuroscience.2013.12.044.

Dumont, V., Giovannella, M., Zuba, D., Clouard, R., Durduran, T., Guillois, B., & Roche-Labarbe, N. (2022). Somatosensory prediction in the premature neonate brain. *Developmental Cognitive Neuroscience, 57*. 101148. https://doi.org/10.1016/j.dcn.2022.101148.

Ellis, A. E., Xiao, N. G., Lee, K., & Oakes, L. M. (2017). Scanning of own- versus other-race faces in infants from racially diverse or homogenous communities. *Developmental Psychobiology, 59*(5), 613–627. https://doi.org/10.1002/dev.21527.

Emberson, L. L. (2017). How does experience shape early development? Considering the role of top-down mechanisms. In J. B. Benson (Ed.). *Advances in child development and behavior* (Vol. 52, pp. 1–41). JAI Press. https://doi.org/10.1016/bs.acdb.2016.10.001.

Emberson, L. L. (2019). How does learning and memory shape perceptual development in infancy? *Psychology of learning and motivation* (Vol. 70, pp. 129–160). Academic Press. https://doi.org/10.1016/bs.plm.2019.03.003.

Emberson, L. L., Richards, J. E., & Aslin, R. N. (2015). Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months. *Proceedings of the National Academy of Sciences of the United States of America, 112*(31), 9585–9590. https://doi.org/10.1073/pnas.1510343112.

Flaten, E., Marshall, S. A., Dittrich, A., & Trainor, L. J. (2022). Evidence for top-down metre perception in infancy as shown by primed neural responses to an ambiguous rhythm. *European Journal of Neuroscience, 55*(8), 2003–2023. https://doi.org/10.1111/ejn.15671.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622.

Gandolfo, M., & Downing, P. E. (2019). Causal evidence for expression of perceptual expectations in category-selective extrastriate regions. *Current Biology, 29*(15), 2496–2500.e3. https://doi.org/10.1016/j.cub.2019.06.024.

Gelskov, S. V., & Kouider, S. (2010). Psychophysical thresholds of face visibility during infancy. *Cognition, 114*(2), 285–292. https://doi.org/10.1016/j.cognition.2009.09.012.

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience, 14*(5), 350–363. https://doi.org/10.1038/nrn3476.

Hunter, B. K., & Markant, J. (2021). Differential sensitivity to species- and race-based information in the development of attention orienting and attention holding face biases in infancy. *Developmental Psychobiology, 63*(3), 461–469. https://doi.org/10.1002/dev.22027.

Hupé, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature, 394*, 784–787. https://doi.org/10.1038/29537.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science, 18*(12), 1084–1089. https://doi.org/10.1111/j.1467-9280.2007.02029.x.

Kok, P., Mostert, P., & de Lange, F. P. (2017). Prior expectations induce prestimulus sensory templates. *Proceedings of the National Academy of Sciences of the United States of America, 114*(39), 10473–10478. https://doi.org/10.1073/pnas.1705652114.

Kouider, S., Long, B., Le Stanc, L., Charron, S., Fievet, A. C., Barbosa, L. S., & Gelskov, S. V. (2015). Neural dynamics of prediction and surprise in infants. *Nature Communications, 6*. 8537. https://doi.org/10.1038/ncomms9537.

Kubicek, C., de Boisferon, A. H., Hillairet de Boisferon, A., Dupierrix, E., Pascalis, O., Lœvenbruck, H., Gervain, J., & Schwarzer, G. (2014). Cross-modal matching of audio-visual German and French fluent speech in infancy. *PLoS One, 9*(2). e89275. https://doi.org/10.1371/journal.pone.0089275.

Lamme, V. A. F. (2018). Challenges for theories of consciousness: Seeing or knowing, the missing ingredient and how to deal with panpsychism. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 373*(1755). 20170344. https://doi.org/10.1098/rstb.2017.0344.

Lamme, V. A., Supèr, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology, 8*(4), 529–535. https://doi.org/10.1016/S0959-4388(98)80042-1.

Liu, S., Quinn, P. C., Wheeler, A., Xiao, N. G., Ge, L., & Lee, K. (2011). Similarity and difference in the processing of same- and other-race faces as revealed by eye tracking in 4- to 9-month-olds. *Journal of Experimental Child Psychology, 108*(1), 180–189. https://doi.org/10.1016/j.jecp.2010.06.008.

Markant, J., Oakes, L. M., & Amso, D. (2016). Visual selective attention biases contribute to the other-race effect among 9-month-old infants. *Developmental Psychobiology, 58*(3), 355–365. https://doi.org/10.1002/dev.21375.

Markant, J., & Scott, L. S. (2018). Attention and perceptual learning interact in the development of the other-race effect. *Current Directions in Psychological Science, 27*(3), 163–169. https://doi.org/10.1177/0963721418769884.

Maurer, D., & Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology, 56*(2), 154–178. https://doi.org/10.1002/dev.21177.

Minar, N. J., & Lewkowicz, D. J. (2018). Overcoming the other-race effect in infancy with multisensory redundancy: 10–12-month-olds discriminate dynamic other-race faces producing speech. *Developmental Science, 21*(4). e12604. https://doi.org/10.1111/desc.12604.

Nakashima, Y., Kanazawa, S., & Yamaguchi, M. K. (2021). Perception of invisible masked objects in early infancy. *Proceedings of the National Academy of Sciences of the United States of America, 118*(27). e2103040118. https://doi.org/10.1073/pnas.2103040118.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences, 11*(12), 520–527. https://doi.org/10.1016/j.tics.2007.09.009.

Prunty, J. E., Jackson, K. C., Keemink, J. R., & Kelly, D. J. (2020). Caucasian infants' attentional orienting to own- and other-race faces. *Brain Sciences, 10*(1). 53 https://www.mdpi.com/2076-3425/10/1/53.

Quinn, P. C., Lee, K., & Pascalis, O. (2019). Face processing in infancy and beyond: The case of social categories. *Annual Review of Psychology, 70*(1), 165–189. https://doi.org/10.1146/annurev-psych-010418-102753.

Quinn, P. C., Lee, K., Pascalis, O., & Tanaka, J. W. (2016). Narrowing in categorical responding to other-race face classes by infants. *Developmental Science, 19*(3), 362–371. https://doi.org/10.1111/desc.12301.

Quinn, P. C., Lee, K., Pascalis, O., & Xiao, N. G. (2020). Emotional expressions reinstate recognition of other-race faces in infants following perceptual narrowing. *Developmental Psychology, 56*(1), 15–27. https://doi.org/10.1037/dev0000858.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*(1), 79–87. https://doi.org/10.1038/4580.

Santolin, C., & Saffran, J. R. (2018). Constraints on statistical learning across species. *Trends in Cognitive Sciences, 22*(1), 52–63. https://doi.org/10.1016/j.tics.2017.10.003.

Sousa, S. S., Amaro, E., Jr., Crego, A., Goncalves, O. F., & Sampaio, A. (2018). Developmental trajectory of the prefrontal cortex: A systematic review of diffusion tensor imaging studies. *Brain Imaging and Behavior, 12*(4), 1197–1210. https://doi.org/10.1007/s11682-017-9761-4.

Squire, R. F., Noudoost, B., Schafer, R. J., & Moore, T. (2013). Prefrontal contributions to visual selective attention. *Annual Review of Neuroscience, 36*, 451–466. https://doi.org/10.1146/annurev-neuro-062111-150439.

Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience, 15*(11), 745–756. https://doi.org/10.1038/nrn3838.

Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., & Hirsch, J. (2006). Predictive codes for forthcoming perception in the frontal cortex. *Science, 314*, 1311–1314.

Vogel, M., Monesson, A., & Scott, L. S. (2012). Building biases in infancy: The influence of race on face and voice emotion matching. *Developmental Science, 15*(3), 359–372. https://doi.org/10.1111/j.1467-7687.2012.01138.x.

Werchan, D. M., Collins, A. G. E., Frank, M. J., & Amso, D. (2016). Role of prefrontal cortex in learning and generalizing hierarchical rules in 8-month-old infants. *Journal of Neuroscience, 36*(40), 10314–10322. https://doi.org/10.1523/jneurosci.1351-16.2016.

Wheeler, A., Anzures, G., Quinn, P. C., Pascalis, O., Omrin, D. S., & Lee, K. (2011). Caucasian infants scan own- and other-race faces differently. *PLoS One, 6*(4). e18621. https://doi.org/10.1371/journal.pone.0018621.

Xiao, N. G., & Emberson, L. L. (2019). Infants use knowledge of emotions to augment face perception: Evidence of top-down modulation of perception early in life. *Cognition, 193*. 104019. https://doi.org/10.1016/J.COGNITION.2019.104019.

Xiao, N. G., & Emberson, L. L. (2023). Visual perception is highly flexible and context dependent in young infants: A case of top-down-modulated motion perception. *Psychological Science, 34*(8), 875–886. https://doi.org/10.1177/09567976231177968.
Xiao, N. G., Wu, R., Quinn, P. C., Liu, S., Tummeltshammer, K. S., Kirkham, N. Z., Ge, L., Pascalis, O., & Lee, K. (2018). Infants rely more on gaze cues from own-race than other-race adults for learning under uncertainty. *Child Development, 89*(3), e229–e244. https://doi.org/10.1111/cdev.12798.