Vision Research 184 (2021) 1–7

Contents lists available at ScienceDirect

Vision Research

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

Development of infants' representation of female and male faces

Scott P. Johnson [a,*,1], Nicholas P. Alt [b], Chibuzor Biosah [c], Mingfei Dong [d], Brianna M. Goodale [a], Damla Senturk [d], Kerri L. Johnson [a,e]

[a] Department of Psychology, UCLA, United States
[b] Department of Psychology, California State University, Long Beach, United States
[c] Program on Ethics, Politics, and Economics, Yale University, United States
[d] Department of Biostatistics, UCLA, United States
[e] Department of Communication, UCLA, United States

ARTICLE INFO

Keywords:
Face preference
Gender-based categorization
Infant perception
Visual development

ABSTRACT

We examined development of 5- and 10.5-month-old infants' face representations, focusing on infants' discrimination and categorization of female and male faces. We tested for gender-based preferences and categorization of female and male faces by presenting infants with pairs of faces and then habituating them to a series of majority female or male face ensembles. We then tested for gender preferences with new face pairs (one female and one male; Study 1) or new face ensembles (majority female and majority male; Study 2). We found that both 5- and 10.5-month-old infants discriminated female from male faces in face pairs, and both age groups looked more at female faces during habituation. Neither age group, however, provided evidence of gender-based categorization. We interpret these findings within a theoretical framework that stresses environmental exposure to different social categories, and infants' ability to detect commonalities of features within categories. We conclude that infants' gender-based categorization of faces is constrained by the set of features available in the input.

1. Introduction

Early in postnatal development, the visual system becomes calibrated to the visual environment—that is, perception becomes attuned to distinctions in available features in visual stimuli (Gibson, 1969; Johnson, 2010). A prominent example of visual calibration comes from studies of face perception in infancy, in particular the means by which infants come to recognize distinctions in social categories such as gender. Faces are among the most important visual stimuli in our environment and the adult visual system is expert at recognizing individual faces and classifying them into distinct social categories (Hancock, Bruce, & Burton, 2000; Haxby, Hoffman, & Gobbini, 2000; Jacques & Rossion, 2006). Young infants show sensitivity to the social category of gender, evinced by a reliable visual preference for a female vs. male face from at least 3–4 months of age when the two faces are viewed side-by-side (Quinn, Yahr, Kuhn, Slater, & Pascalis, 2002). The female face preference, and development of face processing overall, are strongly influenced by visual experience (e.g., Le Grand, Mondloch, Maurer, & Brent, 2001; Macchi Cassia, Kuefner, Picozzi, & Vescovo, 2009; Pascalis et al., 2005; Kelly et al., 2007): Infants generally accrue more experience with female than with male faces because they spend most of their waking hours with the mother and other female caregivers, according to parental reports (Rennels & Davis, 2008) and images from head-mounted cameras placed on infants (Sugden, Mohamad-Ali, & Moulson, 2014). This experience viewing female faces is thought to lead to a processing advantage for female facial features.

The processing advantage conferred by asymmetrical experience with females vs. males in the infant's social environment may lead to different initial representations for female and male faces. Infants' initial representation of human faces may be more female-like in general because it is based primarily on exposure to female facial features (Ramsey, Langlois, & Marti, 2005), and greater experience processing female features leads to better recognition of individual females than males. Furthermore, men's facial features and the spatial configurations of these features may be more variable than women's features (Hopper, Finklea, Winkielman, & Huber, 2014; Johnston, Kanazawa, Kato, &

* Corresponding author.
E-mail address: scott.johnson@ucla.edu (S.P. Johnson).
[1] Scott P. Johnson, https://orcid.org/0000-0003-2970-184X.

https://doi.org/10.1016/j.visres.2021.02.005
Received 31 July 2020; Received in revised form 27 January 2021; Accepted 15 February 2021
Available online 23 March 2021
0042-6989/

Oda, 1997), making male faces relatively more difficult to recognize on the basis of feature overlap.

The asymmetry in exposure to females and males may also lead to an asymmetry in *categorizing* female and male faces, such that female faces are easier to categorize, at least initially. Young infants can form perceptual categories when presented with sets of stimuli from the same class; by definition individual stimuli within a class will have some features in common, and will also share feature distributions. Categorization is defined as recognizing an unfamiliar stimulus as either part of a familiar (learned) category, or different from that category, when individual category members have been shown to be discriminable (Mareschal, French, & Quinn, 2000). Perceptual categories allow infants to organize their perceptual experiences into groupings that may also come to have conceptual significance for children and adults. By 3–4 months of age, for example, infants have been shown to categorize a range of stimuli including dot patterns as well as real-world images of animals and furniture (Bomba & Siqueland, 1983; Madole & Takes, 1999; Mareschal & Quinn, 2001), though there is no indication that infants at that age have representations of these categories that contain semantic information (e.g., whether a category item purrs or barks). Thus perceptual category representations may not always have the same characteristics as might be expected from the corresponding adult category representations, such as the ability to readily categorize both female and male faces (Huart, Corneille, & Becquart, 2005; Ito & Urland, 2003; Wiese, Kloth, Güllmar, Reichenbach, & Schweinberger, 2012).

Quinn et al. (2002), for example, used a categorization method in which 3- to 4-month-olds were presented with eight different female or male faces, shown in pairs during four 15-s familiarization trials. This was followed at test by two pairs of either female or male faces. One of the test faces (either female or male) was from the familiarized gender category and had been seen before, and the other was drawn from the same gender category but had not been seen before. The other pair of faces were from the other gender category and were both new. Infants looked longer at the novel female face after familiarization with the female faces, evidence that they recognized the familiar face (and found it less interesting). Infants familiarized with male faces, in contrast, did not look more at the novel male face. This study implies that infants recognized females as individuals, whereas male faces were identified only at the summary category level and were not fully discriminated as individuals. Quinn et al. (2002) argued that young infants may be experts at female face processing and encode individual exemplars around a summary prototype for female faces. For male faces, however, infants may have more difficulty recognizing different individuals, consistent with a more novice face representation. A representation of the female prototype face, therefore, emerges earlier in development than a representation of the male prototype face (Ramsey et al., 2005). These effects are thought to stem from asymmetrical exposure to female and male faces in early development, as noted previously.

Surprisingly, there is little published work that has examined the development of female and male face category representations in infancy (cf. Anzures, Quinn, Pascalis, Slater, & Lee, 2010; Di Lorenzo, van den Boomen, Kemner, & Junge, 2020; Leinbach & Fagot, 1993; Younger, 1992). Accordingly, the overall aim of the work that we report here was to investigate female and male face representations in 5- and 10.5-month-olds and test predictions drawn from the account of infant face processing just described.

We tested the following predictions. First, we predicted that infants' preferences for faces in the female category will be stronger in 5-month-olds vs. 10.5-month-olds. This is because younger infants likely have less total perceptual experience with male faces and have spent a greater proportion of their visual experience looking at female faces, and the resulting processing advantage for female features is expected to aid recognition of, and preferences for, female faces. Older infants, in contrast, will have gained more experience with male faces and thus are better able to process male features, and this ability to process male-specific features may increase the likelihood that infants will attend more to faces in the male category (cf. Liu et al., 2015; Ramsey et al., 2005). We tested this prediction by presenting infants with pairs of faces, one female and one male, and recording infants' looking times to each. Evidence for a preference for one of the gender categories comes from greater looking times to one face gender (female or male).

Second, we predicted that infants' categorization of female and male faces would be stronger in 10.5-month-olds vs. 5-month-olds. This is because older infants have relatively more experience viewing both female and male faces and thus should be more familiar with facial features that are diagnostic of each category. Younger infants, in contrast, may be unable to categorize faces by gender, or may be able only to categorize female faces (due to greater everyday experience with females). We tested this by habituating infants to a series of face arrays in which a majority (10) were either female or male and the minority (2) were the other gender. Using arrays of faces as stimuli allows us to present multiple examples within each gender category, and using arrays with a gender majority/minority allows us to examine any changes in infants' attention (e.g., from majority to minority gender) as infants gain experience viewing faces during habituation (see Section 2.4 below). Following habituation infants viewed either pairs of new faces, one female and one male (Study 1), or arrays of new faces, one majority female and one majority male (Study 2). Evidence for categorization comes from greater looking times to the non-habituated (i.e., novel) gender category for a single face (Study 1) or face array (Study 2) following habituation. We recorded infants' eye movements as they habituated to the face ensembles so that we could analyze for individual and group differences in attention to the majority vs. minority faces.

To accomplish these goals, infants in both studies were first shown a female and male face side-by-side to establish baseline preferences (Pretest phase), followed by habituation to ensembles of 12 faces with either a female or a male majority (Habituation phase). After habituation, infants in Study 1 were then shown a new pair of female and male faces, and infants in Study 2 were shown two different ensembles of new faces, one with the same majority gender as seen during habituation, and the other with the other gender as the majority (Test phase; see Fig. 1). We reasoned that sensitivity to the majority gender during habituation would be reflected by a shift in gender preference at test.

2. Methods

2.1. Participants

Seventy-three infants (42 females) composed the final sample, 44 5-month-olds (*M* age = 5.0 months, range = 3.5–6.6 months) and 29 10.5-month-olds (*M* age = 10.5 months, range = 9.4–11.8 months). An additional 15 infants were observed but their data were excluded due to excessive fussiness (9 infants), equipment failure (1 infant), experimenter error (1 infant), poor calibration (1 infant), or persistent inattention (3 infants). Infants were randomly assigned to participate in either Study 1 (twenty-six 5-month-olds and fifteen 10.5-month-olds) or Study 2 (eighteen 5-month-olds and fourteen 10.5-month-olds). Infants' racial/ethnic backgrounds, as identified by parents, were as follows: African-American (3 infants), African-American/White (i.e., African-American mother, White father; 4 infants), Asian (7 infants), Asian/White (3 infants), Asian/Hispanic (1 infant), White/African-American (4 infants), White (30 infants), White/Middle Eastern (1 infant), Hispanic/Asian (1 infant), Hispanic (11 infants), Middle Eastern (2 infants), and South Asian/Indian (2 infants). Race/ethnicity was not provided for 4 infants. Infants were recruited from a database of names provided by the county and contacted by telephone and email. Each infant was provided a small gift (a t-shirt or toy) in appreciation for their participation. Infant participants were treated in accordance with University IRB #10-000619 "Brain Mechanisms of Visual Development."

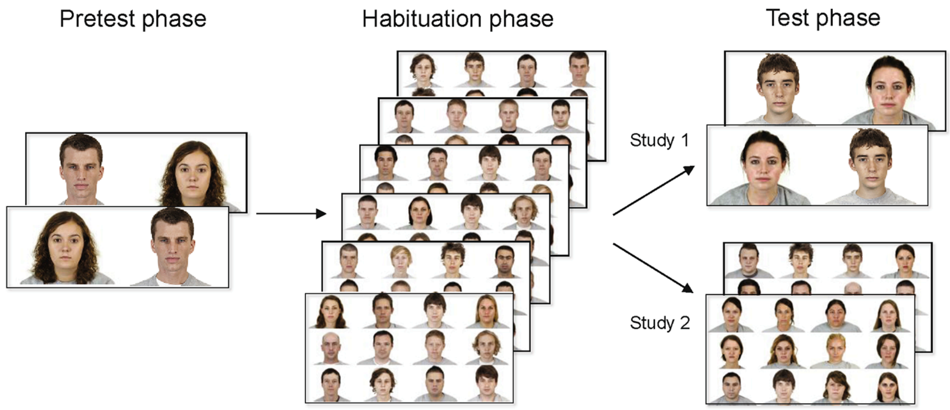

Fig. 1. Schematic depiction of the study design. Pretest phase: Infants viewed pairs of faces, one female and one male (two trials). Habituation phase: Infants were habituated to ensembles of 12 faces, the majority (10) of which were either female or male. Test phase: Infants were shown two new faces, female and male (two trials, Study 1), or two new ensembles, majority female or male (four trials, Study 2).

2.2. Stimuli

Stimuli consisted of a set of 73 White faces (37 female and 36 male) from the Chicago face database (Ma, Correll, & Wittenbrink, 2015; see Fig. 1). To create the stimulus set, Ma et al. recruited adult volunteers between 18 and 40 years of age who were digitally photographed under standardized conditions (e.g., lighting, expression, etc.). Faces were sized to maintain similar dimensions (e.g., eye distance). These photographs were then viewed by a separate group of adult participants who provided subjective ratings of face attributes including masculinity, femininity, attractiveness, and happiness on a 1–5 scale (see Ma et al. for details).

2.3. Procedure

Infants were seated on a parent's lap approximately 60 cm from a computer monitor. A trained observer in an adjacent room viewed the infant on a monitor and coded attention to the screen via button press. The observer's button presses were used to initiate and advance the trials and to code infants' looking times, the principal dependent variable in this study, during the Pretest, Habituation, and Test phases (see below). The observer was blind to the stimulus the infant viewed but was aware of each phase of the study as it progressed.

In addition, we recorded infants' visual attention with an EyeLink 1000 eye tracker (SR Research, Ottawa, ON). Each infant's point of gaze was calibrated prior to testing using a standard 5-point calibration scheme (Gredebäck, Johnson, & von Hofsten, 2009) and points of gaze were recorded at 500 Hz. The gaze data were chiefly used to compute a Majority Preference measure registering infants' relative attention to the majority face gender during the Habituation phase of the study. This was accomplished by summing gaze points within each 4×3 region of the display and noting whether a male or female face was in that region, such that 0.83 represented chance level preference (10/12 faces). As reported below, we analyzed for (a) changes in Majority Preference during habituation as infants learned the face ensembles and (b) relations between Majority Preference during habituation and gender preference at test.

Each study had three phases: Pretest, Habituation, and Test (see Fig. 1). An attention-getter was presented prior to each trial to center the infant's gaze. When the observer determined that the infant was looking at the attention-getter, a button press initiated the next trial and the stimulus was shown.

In the Pretest phase, infants viewed a pair of faces, one female and one male, for two 10-s trials. The observer coded looking to the left or right of the screen by pressing one of two keys on the computer keyboard, and each trial ended when 10 s of looking had accumulated. The left-right positions of the two faces were initially randomized and then switched across trials. The female and male faces seen during the Pretest phase were drawn randomly from the larger set and were not shown again during Habituation or Test phases.

The Habituation phase then commenced immediately and consisted of repeated presentations of 12-face ensembles, either majority (10) female or male. Positions of female and male faces were determined randomly for each array. Faces were drawn randomly from the larger set with the constraints that faces seen during habituation were not shown during Pretest or Test phases. Each ensemble was presented until the infant looked away for more than 2 s (if attention was directed back at the stimulus, the trial continued) or until 60 s elapsed. Each ensemble presented during habituation was distinct in terms of face composition and placement. We used an infant-controlled procedure wherein ensembles were presented until the infant habituated according to a criterion (viz., accumulated looking across 4 trials less than half of the first 4 trials) or looked for 12 trials. The observer coded total looking toward the screen by pressing and releasing a key on the computer keyboard. As noted previously, visual attention to individual female and male faces was recorded at the same time by an eye tracker.

The Test phase then followed immediately and consisted either of a new pair of faces, one female and one male, for two 10-s trials (Study 1) or two new face ensembles, each presented for two trials in alternation (four trials total) that lasted until infants looked away or 60 s had elapsed (Study 2). The observer coded looking to the left or right of the screen (Study 1) or total looking at the screen (Study 2). Side of presentation (Study 1) and presentation order (Study 2) were counterbalanced.

2.4. Statistical analysis

We report two sets of analysis. The first set analyzed infants' looking during Pretest and Habituation phases of Studies 1 and 2 to elucidate preferences for faces in one of the gender categories (e.g., a preference for female) and shifts in preferences over time. These analyses included data from both studies to increase statistical power, as procedures during Pretest and Habituation phases were identical. (Preliminary analyses confirmed that there were no reliable differences between studies in these measures, $ps > 0.20$.) Repeated measures ANOVA on Trial and Stimulus was used to model the infants' looking times (from the observer's button press times) to faces during the Pretest phase and the Test phases of Studies 1 and 2. As noted previously, gaze data from the Habituation phase were modeled as Majority Preference, defined as the ratio of accumulated gaze points (from the eye tracker) toward majority faces (e.g., 10 female vs. 2 male) to the total gaze toward majority and minority (all 12) faces, representing infants' preference for majority faces. The eye tracker recorded accumulated gaze points within "areas of interest" surrounding each individual face (female or male) in the 12-item array. To model the longitudinal trajectories of infants' looking, we used generalized linear mixed models (GLMM) with main effects of time, habituation condition (majority female or male ensembles), and age (5 and 10.5 months) and subject level random intercepts and slopes. Age was not found significant for Majority Preference and was removed from the final model for this outcome. Preliminary analysis showed that the total number of trials until habituation (ranging between 5 and 12) did not have significant correlations with other variables. Therefore we normalized the number of habituation trials to the (0,1) interval for each subject to unify analysis, with zero and one representing the trials from the beginning and end of habituation, respectively, and with other trials spread out equidistantly across habituation.

GLMMs account for correlations between repeated measures within subjects, easily allow for both fixed and time-varying covariates, and automatically handle missing data, thereby producing unbiased estimates as long as observations are missing at random. Seven out of seventy-three infants with no eye tracking data were excluded in modeling the Habituation phase trends. Data on a total of 23 trials (from 10 infants) with zero gaze data for majority or minority faces were considered missing in modeling Majority Preference due to recording error. Effect sizes were reported using Cohen's f based on the mixed effects model where effects of 0.10, 0.25, and 0.40 are generally regarded as small, moderate, and large, respectively.

The second set of analyses related Habituation phase trends to Test phase outcomes (viz., posthabituation novelty preference). Subject-specific intercepts (representing average outcome value during habituation) and slopes (representing rate of change in outcomes during habituation) were obtained based on GLMMs on centered time. Subject-specific intercepts and slopes from the Habituation phase were included among the predictors to model Test phase gender preference, defined as the ratio of looking times (from the observer's button press times) to females over total looking times at females and males, averaged across trials. Other variables in the regression included age (5 and 10.5 months) and habituation condition. We also modeled change in gender preference from Pretest to Test phases.

3. Results

3.1. Pretest phase

We first analyzed looking times (from the observer's button press times) to the female and male face during the Pretest phase. A 2 (trial) × 2 (stimulus: female vs. male) × 2 (infant gender) × 2 (age group) ANOVA, with repeated measures on the first two variables, revealed a significant effect of stimulus, $F(1, 69) = 4.70$, $p = .034$, $\eta^2_p = 0.064$, which was qualified by higher-order interactions between stimulus and age, $F(1, 69) = 4.85$, $p = .031$, $\eta^2_p = 0.066$, and trial, stimulus, and age, $F(1, 69) = 9.48$, $p = .003$, $\eta^2_p = 0.121$ (see Fig. 2). We conducted separate ANOVAs by age to interpret these interactions. A trial × stimulus ANOVA on looking times from the 5-month-old age group yielded a significant main effect of stimulus, $F(1, 43) = 9.05$, $p = .004$, $\eta^2_p = 0.174$, reflecting an overall female preference. Tests for simple effects revealed no significant preference on the first trial, $F(1, 43) = 0.13$, $p = .72$, $\eta^2_p = 0.003$, and a reliable female preference on the second trial, $F(1, 43) = 13.18$, $p = .001$, $\eta^2_p = 0.235$. A trial × stimulus ANOVA on looking times from the 10.5-month-old age group, in contrast, yielded a significant trial × stimulus interaction, $F(1, 28) = 7.86$, $p = .009$, $\eta^2_p = 0.219$ and no other significant effects. Tests for simple effects revealed a significant female preference on the first trial, $F(1, 28) = 6.63$, $p = .016$, $\eta^2_p = 0.192$, and a reliable *male* preference on the second trial, $F(1, 28) = 4.79$, $p = .037$, $\eta^2_p = 0.146$. (Patterns of gaze to the two faces, recorded by the eye tracker, corroborated these results, but are not reported here.)

Next, we tested for the possibility that performance may stem from preferences for particular face characteristics, in particular femininity, masculinity, attractiveness, and happiness ratings provided by adults (Ma et al., 2015). The female faces viewed by infants were rated as more feminine ($t(72) = 25.39$, $p < .001$) and less masculine ($t(72) = -25.21$, $p < .001$) than male faces, but there were no significant differences in attractiveness ($t(72) = 0.93$, $p = .36$) or happiness ($t(72) = 0.45$, $p = .653$). Infants' looking times to female and male faces were converted to preference scores (looking to female faces divided by total looking to both faces; female preference $M = 0.54$, $SD = 0.14$) and correlated with femininity, masculinity, attractiveness, and happiness ratings for each face. None of the correlations was statistically significant, $ps > 0.196$.

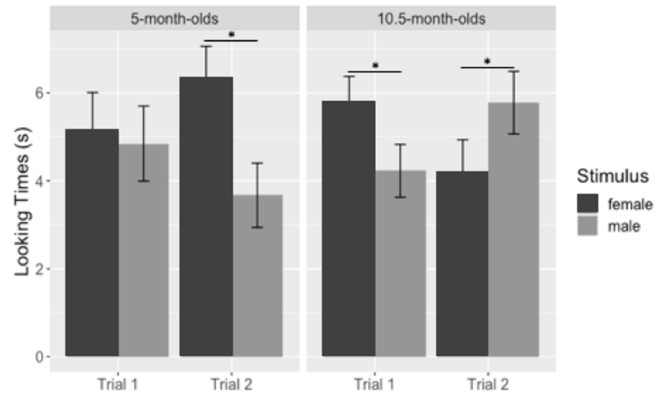

Fig. 2. Looking time results to female and male faces from the Pretest phase. Error bars denote *SEM*.

Finally we tested for the possibility that the female preference we observed (in White face pairs) would be influenced by the individual infants' racial/ethnic background, given that past research has found that the female preference can shift depending on face race (Kim, Johnson, & Johnson, 2015) and the race of the infant (Liu et al., 2015). We compared female face preference scores for infants with White mothers ($N = 35$, M female preference $= 0.57$, $SD = 0.15$) vs. non-White mothers ($N = 34$, M female preference $= 0.52$, $SD = 0.14$) and found no significant difference between them, $t(67) = 1.39$, $p = .168$. (Data from four infants with missing race/ethnicity information were excluded from this analysis.)

In summary, the 5-month-olds we observed exhibited no consistent preference on an initial 10-s trial and showed a female preference on the second trial. The 10.5-month-olds, in contrast, looked longer at the female face on the first trial but reversed preference on the second trial, looking longer at the male face. Infants therefore discriminated faces on the basis of gender, but their preferences did not appear to be driven by differences in the faces' femininity, masculinity, attractiveness, or happiness (as judged by adults), or by individual infants' race/ethnicity.

3.2. Habituation phase

We next modeled trends in Majority Preference during the Habituation phase (from gaze data to majority and minority faces) using main effects of time, habituation condition (habituated to female or to male majority) and age group with subject level random intercepts and slopes using GLMMs. As noted previously, the effect of age group was found not significant and was hence removed from the final model, and none of the effects we report interacted with age. The effect of time was also found not significant, $F(1, 69) = 0.542$, $p = .464$, indicating that the majority preference did not change reliably across habituation. The effect of habituation condition was significant, $F(1, 233) = 31.73$, $p < .001$, Cohen's $f = 0.369$. This effect stemmed from greater accumulated gaze to the majority faces when habituated to majority female face arrays (mean Majority Preference = 0.861) than to majority male face arrays (mean Majority Preference = 0.758; see Fig. 3).

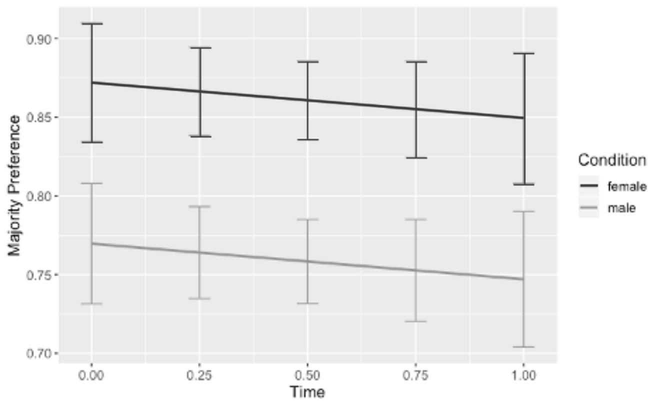

Fig. 3. Gaze data results to majority faces from the Habituation phase. Error bars denote *SEM*.

3.3. Test phase

Looking times (from the observer's button press times) to pairs (Study 1) or arrays (Study 2) of faces during the Test phase were initially modeled with repeated measures ANOVAs on trial, stimulus, habituation condition, and age group, similar to modeling of the Pretest phase looking times. Because the additional factor of habituation condition was added and roughly half the sample size is available in the Test phase (split between Studies 1 and 2), infant gender was removed from the models to reduce model parameters.

In Study 1, a 2 (trial) × 2 (stimulus: female vs. male) × 2 (age group) × 2 (habituation condition: majority female vs. male) ANOVA on post-habituation looking times to female and male faces revealed a main effect of stimulus, $F(1, 37) = 10.15$, $p = .003$, $\eta^2_p = 0.36$, qualified by a higher-order interaction between stimulus and age group, $F(1, 37) = 12.08$, $p = .001$, $\eta^2_p = 0.22$. To interpret the interaction, we conducted separate ANOVAs within each age group. For the 5-month-old age group, a trial × stimulus × condition ANOVA yielded a reliable main effect of stimulus, $F(1, 37) = 26.45$, $p < .001$, $\eta^2_p = 0.52$, demonstrating an overall female preference. To examine the effect of habituation condition on infants' gender preference, we further divided the subjects by condition and conducted separate ANOVAs. The trial × stimulus ANOVA revealed a significant stimulus main effect in 5-month-olds habituated to female face ensembles, $F(1, 12) = 11.91$, $p = .005$, $\eta^2_p = 0.50$, and also in those habituated to male faces, $F(1, 12) = 15.57$, $p = .002$, $\eta^2_p = 0.56$, demonstrating a female preference regardless of the habituation condition (see Fig. 4). Tests for simple effects showed a significant female preference in the 2nd trial of test phase for the 5-month-olds habituated to females, $F(1, 12) = 22.49$, $p < .001$, $\eta^2_p = 0.65$, and in the 1st trial of those habituated to male, $F(1, 12) = 8.075$, $p = .015$, $\eta^2_p = 0.40$. For the 10.5-month-old age group, in contrast, neither the 3-way trial × stimulus × condition nor the 2-way trial × stimulus ANOVA yielded significant effects (Fig. 4).

In Study 2, a 2 (trial) × 2 (stimulus: female vs. male) × 2 (age group) × 2 (habituation condition: majority of female vs. male) ANOVA on posthabituation looking times to majority female and majority male face arrays yielded a trending interaction among trial, stimulus, condition and age group, $F(1, 28) = 3.85$, $p = .060$, $\eta^2_p = 0.12$. As in Study 1, we conducted separate ANOVAs within each age group and then further separated groups by habituation condition. A trial × stimulus × condition ANOVA on looking times from the 5-month-old age group yielded a significant trial × stimulus × condition interaction, $F(1, 16) = 5.17$, $p = 0.037$, $\eta^2_p = 0.24$, indicating significantly different habituation effects between the two trials (see Fig. 5). Five-month-olds tended to look longer at faces of the same gender as the habituation condition in the first trial, but not the second trial. However, there were no significant simple effects of stimulus within each trial. ANOVAs and tests for simple effects of stimulus within each trial on looking times from the 10.5-month-old age group yielded no significant effects (Fig. 5).

3.4. Relating subject-specific trends from the habituation phase to test phase outcomes

To relate Habituation phase trends to Test phase outcomes, we regressed gender preference from the Test phase in Studies 1 and 2, respectively, on age (5 vs. 10.5 months), habituation condition, and subject-specific intercepts (representing average outcome values of Majority Preference) and slopes (representing rate of change in Majority Preference during habituation). In models regressing gender preference on Majority Preference trends, the only significant predictor in the model was found to be age, $F(1, 31) = 10.022$, $p = .003$, Cohen's $f = 0.569$, with a higher preference to female faces in the younger infants, consistent with results of the repeated measures ANOVA on looking times from the test phase (Fig. 5). No term in the regression modeling was found significant in models for a shift in gender preference from pretest to test phase in either study.

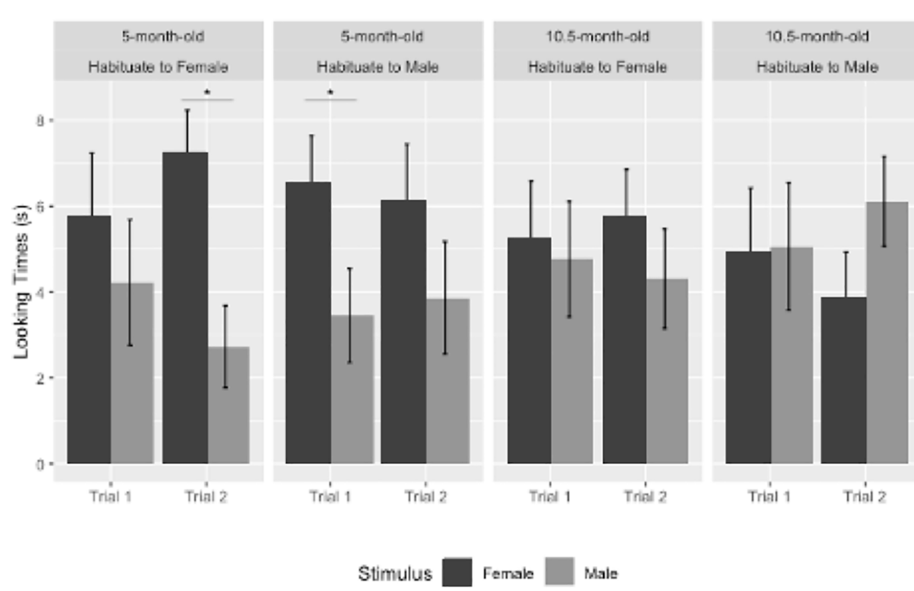

Fig. 4. Looking time results to female and male faces from the Test phase in Study 1. Error bars denote *SEM*.

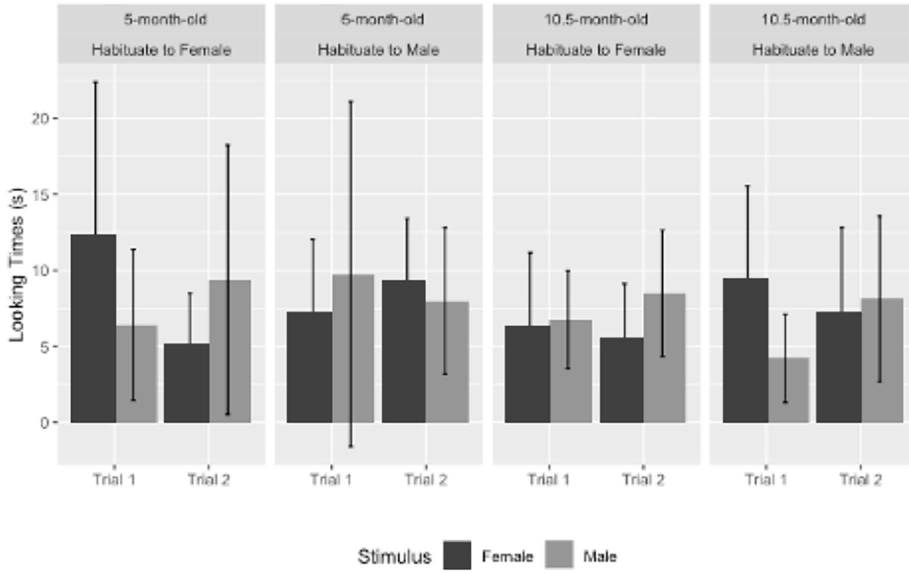

Fig. 5. Looking time results to female and male face arrays from the Test phase in Study 2. Error bars denote *SEM*.

4. Discussion

The present study examined development of infants' gender-based face representations by testing 5- and 10.5-month-olds' preferences for and categorization of female vs. male faces. We found differences in spontaneous gender-based face preferences between the 5- and 10.5-month-old infants. Younger infants showed no preference on an initial 10-s trial, which was followed by a preference for female faces on the second trial. Older infants, in contrast, exhibited a female preference on the first trial, switching to a male preference on the second trial (Fig. 2). However, both age groups showed a stronger preference for majority female faces than majority male preferences during habituation (Fig. 3), and this preference remained consistent across the habituation period. We found no evidence that infants in either age group could categorize female and male faces under tested conditions—that is, there was no evidence of a shift in preference for face gender after habituation to majority female or male faces.

These data provide partial support for the "ladies-first" account of face processing described previously (Ramsey et al., 2005). According to this account, infants' initial face representation is heavily weighted toward female facial features because of an imbalance in exposure to female vs. male faces in the social environment—infants typically are exposed to women more than to men in their everyday life (Rennels & Davis, 2008; Sugden et al., 2014). With development, infants' face representations become more elaborate and come to incorporate male facial features as infants receive more everyday exposure to males. Age differences in spontaneous gender-based face preferences from the Pretest phase are consistent with this account. Five-month-old infants initially showed no preferences, but on a second trial they looked longer at female vs. male faces. 10.5-month-olds, however, showed an initial female preference, switching to a male preference on the second trial. Younger infants, therefore, may have been unable to consistently detect differences between the female and male faces, but with continued exposure, identified gender-specific facial features and looked longer at the female faces, presumably as a result of greater familiarity with female features. Older infants were able to recognize the difference in the two faces more quickly, looking longer at the female face on the first trial and switching to a male preference on the second trial, perhaps due to the relative novelty of male faces (and their characteristic features) in infants' everyday environments. During habituation, both age groups looked relatively more at female faces, which is also consistent with the "ladies-first" hypothesis: The majority preference (i.e., accumulated gaze toward the majority gender in arrays of 12 faces) was significantly higher when infants were habituated to majority females vs. majority males. Unlike age differences in the spontaneous gender-based preferences seen in the Pretest phase, however, the female preference observed during the Habituation phase did not vary by age.

We found no support for the possibility that infants might categorize faces by gender. Though infants clearly discriminated female and male faces in the 12-face arrays (evinced by the stronger majority preference when habituated to majority female ensembles), they did not appear to form a category of either gender that included new items from the same category but excluded items from the other category. As noted previously, infants at 3–4 months have been found to categorize female faces (Quinn et al., 2002), and 3- and 6-month-olds form prototypes of female faces that aid recognition of individuals within the category (de Haan, Johnson, Maurer, & Perrett, 2001; cf. Rubenstein, Kalakanis, & Langlois, 1999). Leinbach and Fagot (1993) reported that 9- and 12-month-olds, but not younger infants, could categorize both female and male faces when the faces were presented individually in a context of gender-typical hairstyles and clothing. It is possible that method or stimulus differences across studies may account for the discrepancy in findings. In the Quinn et al. categorization study, stimuli comprised pictures of female and male models (8 each) with a neutral to positive expression, taken from a catalog. In the Leinbach and Fagot categorization study, likewise, stimuli were pictures of female and male models (12 each) from magazines and catalogs; faces were described as "attractive" and "highly stereotypic as to sex-typical dress and grooming" (p. 320). In the present study, stimuli consisted of faces that were very diverse in appearance (but perhaps even less diverse than real-world faces, which would be inclusive of non-binary individuals), and indeed the adult ratings of these stimuli revealed a substantial range in the faces' attractiveness, prototypicality, and other attributes. Moreover, infants were exposed to at least 36 unique female or male faces across the experiment. The variety of faces, therefore, might have made it difficult to extract commonalities across their features sufficient to form a gender-based category.

5. Concluding remarks

The developments in infants' face gender preferences we report here are a newly-discovered instance of visual calibration to the social environment. Past studies of infants' face preference reported that young infants generally seem to prefer female faces, and our results from

5-month-olds corroborate these findings. This is likely because infants are better able to recognize female vs. male facial features due to greater exposure to women vs. men in real life. As infants gain experience, they come to better recognize men's unique facial characteristics, and thus "lose" the tendency to prefer females, either because female features become more familiar (and thus less interesting), because male features become more familiar (and thus easier to process), or both. However, we obtained no evidence that these changes in visual calibration over the first year after birth yield improvements in categorization of female or male faces. This might be due, at least in part, to our use of face stimuli designed to represent real-world variability in human faces, including female and male face categories (Ma et al., 2015). It is clear that infants were able to discriminate the female and male faces in this stimulus set, but nevertheless were unable to form unique categories of either female or male that excluded individual members of the other category. It remains for future research to determine whether face gender categorization under tested conditions is unavailable until infants gain more experience, or if categorization might be possible with a different set of face stimuli, for example more stereotypically female and male faces (cf. Leinbach & Fagot, 1993).

Funding

This research was supported by NIH grants R01-HD082844 and R01-MH122428. The authors would like to thank the members of the UCLA Babylab for assistance with testing, and especially the infants and their parents for participating. Correspondence can be addressed to Scott P. Johnson, Department of Psychology, UCLA, Los Angeles CA 90095, scott.johnson@ucla.edu.

CRediT authorship contribution statement

Scott P. Johnson: Conceptualization, Methodology, Formal analysis, Project administration, Funding acquisition. **Nicholas P. Alt:** Resources. **Chibuzor Biosah:** Investigation. **Mingfei Dong:** Formal analysis, Visualization. **Brianna M. Goodale:** Resources. **Damla Senturk:** Formal analysis, Visualization, Funding acquisition. **Kerri L. Johnson:** Conceptualization, Methodology, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., & Lee, K. (2010). Categorization, categorical perception, and asymmetry in infants' representation of face race. *Developmental Science, 13*, 553–564. https://doi.org/10.1111/j.1467-7687.2009.00900.x.

Bomba, P. C., & Siqueland, E. R. (1983). The nature and structure of infant form categories. *Journal of Experimental Child Psychology, 35*, 294–328. https://doi.org/10.1016/0022-0965(83)90085-1.

de Haan, M., Johnson, M. H., Maurer, D., & Perrett, D. I. (2001). Recognition of individual faces and average face prototypes by 1- and 3-month-old infants. *Cognitive Development, 16*, 659–678. https://doi.org/10.1016/S0885-2014(01)00051-X.

Di Lorenzo, R., van den Boomen, C., Kemner, C., & Junge, C. (2020). Charting development of ERP components on face-categorization: Results from a large longitudinal sample of infants. *Developmental Cognitive Neuroscience, 45*, 10080. https://doi.org/10.1016/j.dcn.2020.100840.

Gibson, E. J. (1969). *Principles of perceptual learning and development*. Appleton-Century-Crofts.

Gredebäck, G., Johnson, S., & von Hofsten, C. (2009). Eye tracking in infancy research. *Developmental Neuropsychology, 35*, 1–19. https://doi.org/10.1080/87565640903325758.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences, 4*, 330–337. https://doi.org/10.1016/S1364-6613(00)01519-9.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face processing. *Trends in Cognitive Sciences, 4*, 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0.

Hopper, W. J., Finklea, K. M., Winkielman, P., & Huber, D. E. (2014). Measuring sexual dimorphism with a race-gender face space. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 1779–1788. https://doi.org/10.1037/a0037743.

Huart, J., Corneille, O., & Becquart, E. (2005). Face-based categorization, context-based categorization, and distortions in the recollection of gender ambiguous faces. *Journal of Experimental Social Psychology, 6*, 598–608. https://doi.org/10.1016/j.jesp.2004.10.007.

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology, 85*, 616–626. https://doi.org/10.1037/0022-3514.85.4.616.

Jacques, C., & Rossion, B. (2006). The speed of individual face categorization. *Psychological Science, 17*, 485–492. https://doi.org/10.1111/j.1467-9280.2006.01733.x.

Johnson, S. P. (2010). How infants learn about the visual world. *Cognitive Science, 34*, 1158–1184. https://doi.org/10.1111/j.1551-6709.2010.01127.x.

Johnston, R. A., Kanazawa, M., Kato, T., & Oda, M. (1997). Exploring the structure of multidimensional face-space: The effects of age and gender. *Visual Cognition, 4*, 39–57. https://doi.org/10.1080/713756750.

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

Kim, H. I., Johnson, K. L., & Johnson, S. P. (2015). Gendered race: Are infants' face preferences guided by intersectionality of sex and race? *Frontiers in Psychology, 6*, 1330. https://doi.org/10.3389/fpsyg.2015.01330.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Nature, 410*. https://doi.org/10.1038/35073749, 890–890.

Leinbach, M. D., & Fagot, B. I. (1993). Categorical habituation to male and female faces: Gender schematic processing in infancy. *Infant Behavior and Development, 16*, 317–332.

Liu, S., Xiao, N. G., Quinn, P. C., Zhu, D., Ge, L., Pascalis, O., & Lee, K. (2015). Asian infants show preference for own-race but not other-race female faces: The role of infant caregiving arrangements. *Frontiers in Psychology, 6*, 593. https://doi.org/10.3389/fpsyg.2015.00593.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research, 47*, 1122–1135. https://doi.org/10.3758/s13428-014-0532-5.

Cassia, V. M., Kuefner, D., Picozzi, M., & Vescovo, E. (2009). Early experience predicts later plasticity for face processing. *Psychological Science, 20*, 853–859. https://doi.org/10.1111/j.1467-9280.2009.02376.x.

Madole, K. L., & Takes, L. M. (1999). Making sense of infant categorization: Stable processes and changing representations. *Developmental Review, 19*, 263–296. https://doi.org/10.1006/drev.1998.0481.

Mareschal, D., French, R. M., & Quinn, P. C. (2000). A connectionist account of asymmetric category learning in early infancy. *Developmental Psychology, 36*, 635–645. https://doi.org/10.1037/0012-1649.36.5.635.

Mareschal, D., & Quinn, P. C. (2001). Categorization in infancy. *Trends in Cognitive Sciences, 5*, 443–450. https://doi.org/10.1016/S1364-6613(00)01752-6.

Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., & Nelson, C. A. (2005). Plasticity of face processing in infancy. *Proceedings of the National Academy of Sciences of United States of America, 102*, 5297–5300. https://doi.org/10.1073/pnas.0406627102.

Quinn, P. C, Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception, 31*, 1109–1121. https://doi.org/10.1068/p3331.

Ramsey, J. L., Langlois, J. H., & Marti, N. C. (2005). Infants' categorization of faces: Ladies first. *Developmental Review, 25*, 212–246. https://doi.org/10.1016/j.dr.2005.01.001.

Rennels, J. L., & Davis, R. E. (2008). Facial experience during the first year. *Infant Behavior and Development, 31*, 665–678. https://doi.org/10.1016/j.infbeh.2008.04.009.

Rubenstein, A. J., Kalakanis, L., & Langlois, J. H. (1999). Infant preferences for attractive faces: A cognitive explanation. *Developmental Psychology, 35*, 848–855. https://doi.org/10.1037/0012-1649.35.3.848.

Sugden, N. A., Mohamad-Ali, M. I., & Moulson, M. C. (2014). I spy with my little eye: Typical, daily exposure to faces documented from a first-person infant perspective. *Developmental Psychobiology, 56*, 249–261. https://doi.org/10.1002/dev.21183.

Wiese, H., Kloth, N., Güllmar, D., Reichenbach, J. R., & Schweinberger, S. R. (2012). Perceiving age and gender in unfamiliar faces: An fMRI study on face categorization. *Brain and Cognition, 78*, 163–168. https://doi.org/10.1016/j.bandc.2011.10.012.

Younger, B. (1992). Developmental change in infant categorization: The perception of correlations among facial features. *Child Development, 63*, 1526–1535. https://doi.org/10.1111/j.1467-8624.1992.tb01712.x.