The impact of face masks on infants' learning of faces: An eye tracking study

Michaela C. DeBolt,
Lisa M. Oakes
University of California, Davis

Abstract

This preregistered study examined how face masks influenced face memory in a North American sample of 6- to 9-month-old infants ($N = 58$) born during the COVID-19 pandemic. Infants' memory was tested using a standard visual paired comparison (VPC) task. We crossed whether or not the faces were masked during familiarization and test, yielding four trial types (masked-familiarization/masked-test, unmasked-familiarization/masked-test, masked-familiarization/unmasked-test, unmasked-familiarization/unmasked-test). Infants showed memory for the faces if the faces were unmasked at test, regardless of whether or not the face was masked during familiarization. However, infants' did not show robust evidence of memory when test faces were masked, regardless of the familiarization condition. In addition, infants' bias for looking at the upper (eye) region was greater for masked than unmasked faces, although this difference was unrelated to memory performance. In summary, although the presence of face masks does appear to influence infants' processing of and memory for faces, they can form memories of masked faces and recognize those familiar faces even when unmasked.

Keywords

Eye tracking; Infant attention; COVID-19; Face masks; Face processing

One consequence of the COVID-19 pandemic is that face masks are part of our everyday experience, raising questions about how the prevalence of face masks influences face processing. Indeed, a recent special issue of the journal *Cognitive Research: Principles and Implications* reported studies that revealed that adults and children are impaired at both *learning* and *recognizing* faces with masks (Carragher and Hancock 2020; Stajduhar et al. 2022). Given the importance of faces for infants' learning, we asked here whether the presence of face masks influenced infants' learning of and memory for specific faces. The effect of face masks may be especially pronounced on infants' face processing, given the dramatic changes that occur in face processing during the first years of postnatal life (Pascalis et al. 2011). However, it is unknown what impact the presence of face masks will have on face processing for infants who have only experienced a world in which they frequently encounter faces covered with masks. This project addressed this question.

Correspondence should be addressed to Michaela C. DeBolt, Center for Mind and Brain, 267 Cousteau Pl., University of California, Davis, CA 95618 (mdebolt@ucdavis.edu).

Infant face processing

Infants' processing of, learning of, and memory for faces has been studied for decades using the visual paired comparison (VPC) test (Pascalis et al. 1998; Pascalis, de Haan, and Nelson 2002; Fagan 1972). In this task, infants first visually inspect an image of a face for a period between 10 and 30 s, and then they view that now-familiar face paired with a novel face. The initial *familiarization phase* is the period during which infants presumably learn the stimulus. When that now-familiar face is paired with a novel face during the *test phase*, infants' memory for the original face is assessed. As a result, infants' performance in this task, as measured by their preference for the novel stimulus during test, reflects multiple memory processes. Research using the VPC task has revealed several facts about infants' face processing. First, face processing reflects infants' experience. Specifically, whereas 6-month-old infants show novelty preferences when familiarized with faces of familiar or novel races (Kelly et al. 2009, 2007), or human or monkey faces (Pascalis, de Haan, and Nelson 2002; Pascalis et al. 2005), 9-month-old infants show novelty preferences only when familiarized with faces of familiar races or human faces. This suggests that experience processing some faces (human faces of a particular race) shapes infants' face processing, making it easier for them to process relatively familiar faces compared to relatively novel faces. Therefore, infants who have experience seeing masked faces may have acquired strategies to process masked faces.

Second, *holistic* face processing appears to emerge in infancy. One hallmark of mature face processing is that it is holistic; rather than perceiving faces as a collection of parts, perceivers combine individual facial features into an "unbroken whole" or a single perceptual representation (Pascalis et al. 2011). As a result, adults and children can better recognize and identify individual face parts (e.g., a specific nose) when presented *within* the context of a whole face compared to when the face part is presented in isolation (Pellicano and Rhodes 2003; J. W. Tanaka and Farah 1993) phenomenon referred to as the *part-whole* effect (Tanaka & Farah, 1993).

During the first year, infants in Western countries begin to process upright and inverted faces differently (Oakes and Ellis 2013) and recognize configural (verus featural) changes in faces (Schwarzer, Zauner, and Jovanovic 2007), two indicators of holistic face processing. Especially relevant for the current study, Turati et al. (2010) found that 3-month-old infants' face memory appears to be influenced by holistic face processing. They habituated infants with just the top half of a face. During test, the 3-month-old infants in this study discriminated the familiar face from the novel face only if the top and the bottom half of the test faces were misaligned, but not if the top and bottom halves of the test faces were aligned. This suggests that infants were processing the aligned faces holistically, and were unable to match the previously learned top half of the faces to the aligned whole seen during test. Thus, the presence of a face mask that covers the lower portion of the face may impair infants' holistic face processing and have cascading effects on their learning or recognition of faces.

Third, research suggests that, in at least some tasks, infants form memories of the *face*, and not the specific *image*. Kelly and colleagues (2007, 2009), for example, demonstrated that

following familiarization with a face depicted in ¾ view, British White infants and Chinese infants showed a memory for that face shown in a frontal view (or vice versa). Similarly, Otsuka and colleagues (2009) found that 3- to 4-month-old Japanese infants familiarized with a smiling face showed a novelty preference when that person was shown paired with a novel person, both with neutral expressions. Thus, infants do not simply remember a specific image or photo of a face, but can recognize the correspondence between two different images of the same face. Infants may also be able to generalize their representation of an unmasked face to a masked face, or vice versa.

Finally, studies involving eye tracking with infants from Western countries have shown that interest in the eye regions of faces changes with infant age, experience, and testing context. Overall, when viewing static images infants show a strong bias to look at the eyes of faces (Oakes and Ellis 2013; Xiao et al. 2015), but that with increased age infants scan faces more broadly, perhaps focusing on the mouth region because it is important in language processing (Pons, Bosch, and Lewkowicz 2015; Lewkowicz and Hansen-Tift 2012). In Western cultures, attention to the eye regions of faces can predict infants' memory for faces (Amso et al. 2010), and may be related to how they process faces during a learning or habituation phase (Bolhuis, Kolling, and Knopf 2015). Indeed, when familiarized with partly occluded faces, newborns learn and remember faces only if the eye region is visible (Gava et al. 2008). Taken together, this research may indicate that face masks will have little effect on infants' face learning and memory as long as they can still see the eyes.

The effect of masks on face processing

Because adults' face processing is holistic, recognizing faces piecemeal should be particularly difficult and face masks may make it more difficult to learn and recognize faces. Prior to the COVID-19 pandemic some studies have shown that the introduction of props that partially occlude faces, such as sunglasses, does impair face processing in adults (Noyes and Jenkins 2019). As the pandemic progressed, it became clear that face masks would be common in many countries, such as the United States, Canada, and European countries, where face coverings had not previously been common. In response to this change, researchers from these countries conducted studies showing that adults' face processing is impaired by the presence of a mask. For example, adults had more difficulty matching familiar and unfamiliar faces if even just one of the two faces was masked (Carragher and Hancock 2020). Occlusion by masks negatively affects face matching, (Bennetts et al. 2022; Noyes et al. 2021). In addition, adults' ability to process and recognize the emotions of faces is poorer when those faces are masked (Grundmann, Epstude, and Scheibe 2021; Grenville and Dwyer 2022).

A similar effect of masks is seen on Western children's face processing. For example, Stajduhar et al. (2022) found that although children between 6 and 14 years of age had little difficulty identifying upright unmasked faces, their identification of upright masked faces was quite poor. School age children's ability to read emotions seems to be impaired by the presence of a face mask (Carbon and Serrano 2021), although children 7 to 13 years of age can make accurate emotion judgments about masked faces (Ruba and Pollak 2020). Thus, children and adults who were not regularly exposed to face masks prior to the pandemic

seem to have difficulty processing faces with face masks. What is yet unknown is whether infants whose only face experience has occurred during the pandemic–when face masks have been common–process masked and unmasked faces differently.

The present study

We used a VPC procedure to determine whether infants' recognition memory for faces is influenced by the presence of face masks. On each VPC trial, infants were familiarized with an image of a woman's face and then tested with that image paired with an image of a novel woman's face. We used a fully crossed 2 by 2 design in which we varied whether or not the women were wearing a mask during familiarization or test (see Figure 1). Infants received one trial of each type. This design allows us to ask not only how masks influence infants' learning and memory for faces, but also how infants generalize their learning about an individual face between a masked and unmasked condition.

Based on decades of research (Colombo, Mitchell, and Horowitz 1988; Fagan and McGrath 1981), we predicted that infants would show significant novel preferences (i.e., above the chance level of .50) when the women were unmasked during both familiarization and test. This condition is essentially a replication of previous work. We did not have directional hypotheses regarding the other three conditions. The presence of facemasks may not influence infants' responding in this task, either because for them facemasks are relatively familiar or because facemasks do not impede the infants' view of the eyes (or both). In this case, infants will respond in the same way to all four tests. However, if facemasks disrupt infants' holistic face processing or their ability to generalize from one image of a face to another, infants' responding in the four conditions may differ.

Our design also allows us to examine the effect of facemasks during *learning* or encoding versus during recognition. If masks influence infants' *encoding* of faces, we may observe differences in infants' recognition memory as a function of whether or not the faces are masked during familiarization (the period when infants encode the familiar stimulus). Alternatively, if masks influence *recognition* of faces, we may observe differences in recognition memory as a function of whether or not the faces are masked during test. Of course, masks may influence *both* encoding and recognition, in which case we may observe differences regardless of when the masks are present (i.e., during familiarization or test).

Because we conducted this experiment using eye tracking, we also could ask exploratory questions about *where* infants look, and how their patterns of looking are related to their learning. Specifically, we examined where infants looked while viewing masked and unmasked faces. We also measured the age of each infant that participated in our study and parents reported on their infant's experience seeing people wearing face masks (i.e., "face mask experience"). We explored whether infants' novelty preferences on the different trial types varied as a function of these variables.

Method

Participants

This study was preregistered on the Open Science Framework (OSF; https://osf.io/ugfd9/). We identified a target sample size of 55 infants based on a power analysis, adopting a sensitivity approach and simulated beta distributed data using methods described in DeBolt et al. (2020) (see preregistration and supplemental materials for more details). A sample of 55 infants would provide 80% power to detect a significant difference from chance with a one-sample *t*-test when the population mean is between .56 and .57, given the typically observed standard deviations. We tested 68 infants between 6 and 9 months of age between 12/29/21 and 3/22/2022, during which there was a State mask mandate in place. We chose this age range because the literature shows that infants within this age range can rapidly form a memory for a face in the VPC procedure (Rose and Feldman 1987).

As specified in our preregistration, we peeked at the data after testing 32 infants to confirm that our procedures were working. We visualized infants' novelty preference scores for each condition and, although we did not evaluate statistical significance at this point, we established that infants' novelty preference scores for the unmasked familiarization/ unmasked test condition appeared to be above chance. We then adopting a sequential data analysis approach (Schott, et al., (2019) and adjusted our final *p*-values to account for this one peek at the data, using the O'Brien-Fleming spending function (DeMets and Lan 1994; Schott, Rhemtulla, and Byers-Heinlein 2019), that allocates a larger proportion of alpha (and therefore statistical power) to later peeks relative to early peeks. Our adjusted *p*-value threshold to determine statistical significance is .046 calculated based on our *smallest* observed sample size[1].

We excluded infants' data from the analyses using preregistered exclusion criteria of failing to provide usable data on more than 2 trials (before other exclusion criteria were applied) either because we could not successfully calibrate the infant ($N = 5$) or other problems with their eye tracking data (e.g., their right eye could not be tracked or the track was frequently lost throughout each trial; the experimenter made this determination at the time of testing, before the data had been processed or analyzed) ($N = 4$). We also excluded 1 infant based on not meeting our criteria for the duration of looking during familiarization and test (see *Data Processing* later).

Our final sample was 58 infants (*M* age = 221.5 days, *SD* = 32.5). There were 31 girls, 26 boys, and 1 infant whose sex was not reported. Infants resided in the rural, urban, and suburban areas surrounding Davis, CA. Although we do not have data about how much time the infants in our study spent with their caregivers, in this community during this period infants likely were spending significant time with their caregivers due to parents being primary caregivers and not working outside of the home, particularly during this time when many daycares had not yet returned to full capacity. The sample was racially and

---

[1]As is described in the *Data Processing* section, not all infants completed all 4 trials, therefore the *N*s vary across the four trial conditions. The smallest sample size across the 4 trial conditions after accounting for the extreme scores is 50; we used this sample size to calculate the final nominal *p*-value. Note that this value differs from the preregistered *p*-value of .048 based on a total sample size of 58 infants.

ethnically diverse; 38 infants were White (8 Hispanic or Latino/a), 5 were Asian (1 Hispanic or Latino/a), 3 were Black (2 Hispanic or Latino/a), 10 were mixed race (2 Hispanic or Latino/a), and 2 infants' race was unreported (1 Hispanic or Latino/a). Families were middle class; virtually all mothers had earned a high school degree (98%) and 79% had earned at least a 4-year college degree.

Infants were recruited through ads on social media (Facebook, Instagram), word of mouth, or from births recorded by the State Office of Vital Records. Families identified through birth records are sent informational flyers with instructions on how to submit their contact information into our volunteer database of potential research participants. When infants reached the appropriate age for this study, we contacted parents and invited them to participate.

Apparatus

Eye movements were captured using an SMI-RED-M eye tracker at a rate of 120 Hz, which was attached to the bottom of a 22-in LCD monitor. A web-camera attached to the top of the monitor recorded the participants' head and body position. The monitor was affixed to an ergo arm so the experimenter could position the eye tracker to optimally locate each infant's eyes in the center of the detection radius of the eye tracking system. A Dell laptop was used to capture eye gaze data from the eye tracker, present the stimuli, and run the experiment. The infant and eye tracker were separated from the experimenter and the other equipment by a large white cloth that divided the room.

Stimuli

The stimuli were photographs of 8 different women, all between 19 and 22 years of age, and of European ($N = 4$; 1 Latina/Hispanic), East Asian (Chinese; $N = 2$), or South Asian (Indian; $N = 2$) descent. Each woman wore a black tee-shirt and was positioned against a light colored wall. Each woman was photographed (closed-mouth smiling) both with and without a face mask (all women wore the same black face mask). Images were cropped and edited to be approximately the same size, and were 15 cm high by 10.5 – 12 cm wide (10 – 11.42 ˚ by 14.25˚ visual angle at a viewing distance of 60 cm).

We created image pairs, in which two images were presented side-by-side, at a center-to-center distance of 15.75 – 16.5 cm (14.95 – 15.66˚ visual angle). Each image was 10.5 cm (10˚ visual angle) from the center of the screen (where the fixation stimulus was presented between trials) (see Figure 1). The familiarization pairs consisted of two identical images and the test pairs consisted of two different images. Examples of the pairs are presented in Figure 1, and additional pairs can be found in the supplemental materials at https://osf.io/ugfd9/. Note that our pairs were constructed such that individual images were both a familiar stimulus for some infants and a novel stimulus for other infants. Thus, although we did not control for luminance, physical salience, or other physical differences in the stimuli, such differences would not have a systematic effect on our results. In addition, as will be explained in more detail in the results section, our models revealed estimated effects of stimuli that were close to zero (see Tables S2, S7, and S8).

Note that the faces in our stimuli were Asian or White, and that we showed them to infants from a diverse racial and ethnic backgrounds. Although our infants are in the age range in which other-race effects have been observed, infants in our community are exposed to a wide range of face types, including many Asian and White faces, and such heterogeneity of face experience seems to reduce infants' responding to own and other races differently (Ellis et al., 2017). Thus, we expected that for our infants Asian and White faces would be responded to similarly. In addition, each of our stimulus pairs involved two women of different races and/or ethnicities, and thus infants should be able to tell the two stimuli apart even if they have any insensitivity to the distinction between two faces of the same (novel) race.

Between trials, we presented short clips from children's videos (e.g., Sesame Street, Disney movies) to keep the infant interested in looking at the screen in general. During trials, we presented a short video of an animated shape (e.g., a set of shapes rotating in the middle of the screen) accompanied by a short sequence of sounds, to reorient the infant to the center of the screen before the pairs of faces were presented.

Procedure

An example video from an infants' session can be found at https://osf.io/ugfd9/, and the videos of all the sessions (for which parents provided permission to share) can be found at databrary (https://nyu.databrary.org/volume/1161). The present study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any assessment or data collection. All procedures involving human subjects in this study were approved by the IRB at the University of California, Davis. After obtaining informed consent, parents and infants were escorted to a sound attenuated testing room, where infants were seated on their parent's lap or infant high-chair, approximately 60 cm from the monitor. To minimize bias, parents wore felt-covered sunglasses to occlude their view of the stimuli during the experimental trials.

The experimenter located the infants' eyes using SMI software iView, adjusting the position of the monitor (using the ergo arm) as necessary, as the infant watched a child-appropriate video. Next, the experimenter initiated the SMIs automatic calibration sequence, which involved presenting visually interesting stimuli (a looming circle) at each of 5 calibration points. A validation procedure was initiated after calibration; an image of a yellow rubber duck accompanied by a chirping sound was presented in the four corners of the screen. Using the feedback from the validation, the experimenter repeated the calibration until the infant's gaze was within approximately 2° of the validation stimulus locations.

Immediately following calibration, each infant received up to 4 trials, each with the following sequence: 1) An animated video clip presented in the center of the screen; 2) When the infant fixated this video for 500 ms, the *familiarization phase* began in which two identical images of the same woman were presented side-by-side for 12 s; 3) The presentation of an attention-getting stimulus (e.g., rotating squares, swirling shapes); once infants fixated this stimulus for 200 ms, the first 5-s *test phase* began in which the now-familiar woman was paired with an image of a new face (i.e., "novel face"); 4) Another

brief attention-getter and once infants had fixated it for 200 ms, a second 5-s *test phase* was presented with the same faces, but with the left-right positions reversed. The presentation of faces during the familiarization and test phases were accompanied by classical music (Vivaldi violin concerto No. 3 in F Major or Mozart piano concerto in C, K. 467) to help maintain infants' interest. The eye-tracking procedure lasted 2–3 min in total.

Infants received one trial in each of the four conditions presented in Figure 1. In the *masked-familiarization/masked-test* condition, the face was masked during familiarization and both faces were masked during test. In the *unmasked- familiarization/masked-test* condition, the face was unmasked during familiarization and both faces were masked during test. In the *masked-familiarization/unmasked-test* condition, the face was masked during familiarization and neither face was masked at test. Finally, in the *unmasked-familiarization/unmasked-test* condition, the face was unmasked during familiarization and both faces were unmasked during test. The order of the four trials was randomly determined by the eye tracking software, Experiment Center, for each infant.

In addition to a random ordering of trial types, the specific faces infants received varied. We created two different collections of faces, and then generated two stimulus sets with each collection of faces by varying which faces were novel and which were familiar. For example, sets 1 and 2 used the same face, but the familiar face in set 1 was the novel face in set 2. Two different face pairs were used for each trial type (e.g., one face pair for sets 1 and 2 and a different face pair for sets 3 and 4) to ensure that infants' responding in a particular condition was not specific to a single pair of faces (see supplemental materials in our OSF repository, https://osf.io/ugfd9/, for a full schematic diagram of our counterbalancing). After the eye tracking procedure, the parent was asked to report their infant's experience with people wearing face masks. Specifically, they responded to the questions "How often does your infant see members of the household wearing a face mask or face covering?" and "How often does your infant see members of the community wearing a face mask or face covering?" For each question, parents indicated the frequency as daily, several times a week, once a week, once every two weeks, very rarely, or never. Data for one infant was not reported.

Data processing

Using BeGaze, the software provided by SMI, for each infant we exported the *net dwell time*, or the duration as reflected by the total number of samples recorded by the eye tracker in each Area of Interest (AOI) on each trial. Because each sample was 8.33 ms, the net dwell time reflects the number of samples times 8.33 ms. Note that we did not filter our data into fixations for two primary reasons. First, as has been shown by several researchers (Hessels et al. 2016; Wass, Forssman, and Leppänen 2014), data quality can influence how fixations are parsed and it is not entirely clear which fixation algorithm is most appropriate for infant data. Thus, filtering data into fixations can introduce systematic errors into the data. Second, the VPC task used here has traditionally been implemented with human observers rather than eye trackers. Human observers are unable to detect fixations, instead recording the overall time that infants' gaze is directed at one stimulus or another. For these reasons, our conclusions are based on the total amount of looking (not on any metric related to fixations),

and total looking will be very similar regardless of whether or not we applied a fixation filter. For each face, we created separate AOIs for the upper and the lower halves (see Figure 2), bisecting the face in two equal halves. The size of the AOIs varied somewhat for different faces[2], but the two AOIs for a given face covered approximately the same area.

As a first step, we examined the distribution of the total amount of looking (to the upper and lower AOIs for both images combined) to determine how often infants' looked less than 10% of any phase (e.g., less than 1200 ms during familiarization or less than 500 ms on each test phase). We found that 6% of trials contained less than 1200 ms of looking during familiarization, 5% of trials contained less than 500 ms of looking during the first test phase, and 5% of trials contained less than 500 ms of looking during the second test phase. We discarded the 23 trials (10% of all the trials) that did not include at least 1200 ms total looking during the familiarization phase and 500 ms total looking during each of the two test phases (as a result, 1 infant was excluded). Note that some trials met more than one exclusion criteria thus there was overlap across the three criteria. We report in our supplemental materials an evaluation of the robustness of our main analyses to variations in trial exclusion criteria and AOI size. In general, the pattern of results did not vary. In the final data, infants looked on average 6865.17 ms (SD = 2622.62, range from 1571.50 to 11353.50 ms) during familiarization, 3267.68 ms (SD = 1017.24) during the first test phase, and 2998.60 ms (SD = 1047.47) during the second test phase.

Next, we used these looking times to calculate our primary measure, the novelty preference score, on each trial from the infants' total looking time (i.e., at the two AOIs combined) to each face on each trial. We divided the amount of time infants' gaze was directed to the novel face by the total amount of time gaze was directed to the novel and familiar faces combined. We calculated a separate novelty preference score for each infant and for each of the four trials (one for each condition).

Analytic approach

Our primary preregistered analyses were separate *t*-tests comparing the novelty preference score for each trial type to chance (.50). As described earlier, we adjusted our nominal *p*-value to account for our initial peek after 32 infants had been tested, and our significance thresholds for our planned analysis were adjusted from .05 to .046; *p*-values from our main analysis must meet this criterion to be considered statistically significant (see https://osf.io/ugfd9/ for more details). In addition, to establish whether novelty preference changed over age, we conducted zero-order correlations between infant age and novelty preference scores for each trial type.

We also present several exploratory analyses examining where infants' looked when viewing masked and unmasked faces, as well as the relation between infants' looking during the familiarization (learning) phase of each trial and their novelty preference during test and age. Our first analysis was a LMM on the total duration of looking during familiarization (i.e.,

---

[2]We also created larger rectangular AOIs that were of equal size for all faces; these AOIs corresponded to the "screen" in a traditional VPC task. The patterns of novelty preference scores derived from the larger AOIs are nearly identical to the results reported in the main text (see Figure S4 and Table S12).

the upper and lower regions combined) with the familiarization condition as a fixed effect. In this model, we included a random intercept of *familiarization stimulus ID* that reflected the 8 unique faces presented during familiarization. We included random intercepts for participant and stimulus. This model was defined as (note, we indicate each model with a unique letter, i.e., "MA" here, to help connect the models to the actual analyses described later):

MA: lmer(Total Looking During Fam.~Fam. . Condition(masked vs . unmasked) + (1│Participant) + (1 | Fam . Stimulus ID))

We next examined whether infants' preference for the upper half of the face differed when the familiarized face was masked compared to when the familiarized face was unmasked. To that end, we fitted a LMM to infants' preference for the upper half of the face with a fixed effect of familiarization condition including the same random intercepts as the previous model. This model was defined as:

MB: lmer(Pref. for Upper Half During Fam.~Fam. Condition(masked vs. unmasked) + (1 | Participant) + (1 | Fam. Stimulus ID))

We also examined how infants' infants' distribution of looking to the upper half of the face *changed* from familiarization to test. We fitted a LMM to infants' proportion of looking to the upper half of the face with trial *phase* (familiarization versus test) and *trial condition* (see trial conditions listed in Table 1) as fixed effects. We also included a 2-way interaction term between phase and trial condition to examine whether infants' preference for the upper half of the face changed from familiarization to test differently across trial types. Again, we included a random intercept for participant and for stimulus ID. The model was defined as:

MC: lmer(Pref for Upper Half~Trial Phase(Fam. vs. test) X Trial Condition(4 levels) + (1│Participant) + (1 │Stimulus ID))

Omnibus *F*-statistics were used to evaluate the significance of the fixed effects and interactions from some of the LMMs (Luke 2017).

Finally, because research has revealed that the proportion of time infants spend looking at the upper regions of faces, or the eye region, compared to the lower region, or mouth region, varies with age (Oakes and Ellis 2013; Ayneto and Sebastian-Galles 2017), we evaluated the correlation between infants' looking to the upper half of the face with age.

We also examined the correlation between infants' novelty preference scores and 1) infants' looking to the upper half of the face during familiarization, and 2) infants' mask experience. These correlations did not yield any significant effects so they will not be reported here. Details can be found in the supplementary materials (see Figure S2 and Tables S5 and S11).

Results

All analysis code and data used to generate the following results can be found in the OSF repository: https://osf.io/ugfd9/. Overall, infants contributed an average of 3.82 trials (SD = .57) to the analyses (See Table 1 for the number of trials completed per condition).

Preregistered analysis: Infants' novelty preference during test

Our first analysis was our primary analysis examining the novelty preference score for each condition compared to chance (.50) with one-sample (two-tailed) *t*-tests. The mean novelty preference score for each trial type (condition) is presented in Figure 3, and the mean scores and comparisons to chance are in Table 1. When infants are tested with masked faces, their novelty preference score is not greater than chance, regardless of whether or not the familiarization face was masked. When the faces during the test phase were not masked, infants had a tendency to prefer the novel face (i.e., both sample means were .55).

Although our preregistered analysis plan did not involve excluding extreme scores from these analyses, inspection of the individual infants' responding in Figure 3 indicated that there were some extreme scores near zero. We examined the novelty preferences scores collapsed across trial conditions using the *boxplot* function in R and identified three extreme scores (i.e., scores that are 1.5 times the interquartile range). We conducted *t*-tests excluding these three scores (see Table 1); when the faces during test were *not* masked, infants' novelty preference scores were greater than chance, regardless of whether or not the familiarization face was masked. Excluding the extreme scores in the masked-familiarization/masked-test condition did not change the outcome of the *t*-test.

These results were confirmed with binomial tests on the proportion of infants in each condition who had scores above chance, excluding the three extreme scores. Although the proportion of infants in both the unmasked-familiarization/masked-test and the masked-familiarization/masked-test conditions with scores above .50 was not greater than 50%, *p*s > .89, the proportion of infants in both the unmasked-familiarization/unmasked-test and the masked-familiarization/unmasked-test conditions with scores above .50 was significantly greater than 50%, *p*s ≤ .03. In summary, therefore, infants appeared to show recognition memory for the face learned during familiarization as long as the test faces were unmasked during test.

We also conducted a number of exploratory analyses evaluating infants' preferences within a Bayesian framework (see supplemental results for details) that provided additional insight into infants' preferences. Specifically, the Bayesian indices suggest that the evidence for the null is *weaker* in the masked-familiarization/masked-test condition relative to unmasked-familiarization/masked-test condition, warranting a more cautious interpretation of infants' memory for masked faces when tested with masked faces. We also conducted exploratory analyses comparing performance on the different trial types using LMMs. These analyses did not reveal any differences (see supplemental analyses).

Preregistered analysis: Correlations with age and preference scores

We tested whether novelty preference scores were associated with infant age by examining the correlation between infants' age in days with their novelty preference score. The scatterplots are presented in Figure 4, and as can be seen, none of the correlations were significant except for the masked familiarization/masked test condition (both including and excluding extreme scores, see Table 2). Overall, therefore, infants' novelty preference was

not related to age, with the exception of when both the familiarization and test faces are masked.

Exploratory analysis: Infants' looking during learning

We asked how infants' looking during learning (i.e., the familiarization phase) differed as a function of mask condition during familiarization. The LMM described in the Analytic Approach section (see model MA) revealed that the infants' total duration of looking during familiarization did not vary as a function of mask condition, $F(1, 154.52) = .83$, $p = .36$ (full details can be found in Table S7 in the supplemental results); the duration of infants' looking did not statistically differ when familiarization with masked faces ($M = 6612.99$, $SE = 314.10$) and unmasked faces ($M = 6843.81$, $SE = 313.00$).

We also calculated the proportion of infants' looking to the upper half on the familiarization and test phases by dividing the amount of time infants' gaze was directed to the upper half of the faces by the total amount of looking directed to the two face images. Comparison of these scores to chance revealed that infants overwhelmingly preferred to look to the upper half of the face across all conditions and phases, all ps < .001 (see Figure 5, and Table S9 in the supplemental results).

Although chance comparisons revealed that infants preferred to look at the upper half of the face overall, a comparison of the points in Figure 5 suggests that infants' preference for the upper half of the face was stronger when the familiarized face was masked compared to when the familiarized face was unmasked. Indeed, the LMM described in the Analytic Approach section (see model MB) revealed an effect of familiarization condition, $\beta = -0.13$, $t = -6.32$, $p < .001$, confirming the observation that when the familiarized face was masked, infants' preference for the upper half of the face was stronger than when it was unmasked (see Table S7 in the supplemental results).

Further understanding is gained by examining how infants' distribution of looking *changed* from familiarization to test. Inspection of Figure 5 shows that the proportion of time that infants devoted looking to the upper half of the face did not change from familiarization to test in the masked-familiarization/masked-test condition and in the unmasked-familiarization/unmasked-test condition, but when infants saw masked faces *only* during familiarization *or* test, the proportion of looking to the upper half appeared to change from familiarization to test. To confirm this observation, we fitted the LMM described in the Analytic Approach section earlier (see model MC). This analysis revealed a significant main effect of trial condition, $F(3, 7.66) = 9.51$, $p = .006$, qualified by a significant interaction between trial condition and phase, $F(3, 349.98) = 14.83$, $p < .001$ (full details can be found in Table S8 in the supplemental materials). This interaction demonstrates that the change in upper half preference from familiarization to test differed in the four trial conditions. When familiarized with a masked face and then tested with unmasked faces, infants' proportion of looking to the upper half of the face decreased from familiarization to test $t(53) = 4.40$, $p < .001$. Similarly, when familiarized with unmasked faces and tested on masked faces, infants' proportion of looking to the upper half of the face increased from familiarization to test , $t(53) = 5.02$, $p < .001$. In other words, in these mixed conditions, infants scanned the faces more broadly when no face mask was present. When there was no change from

familiarization to test with respect to masks, the proportion of infants' looking to the upper half of the face did not change, $ps \geq .25$. In general, therefore, infants looked at masked and unmasked faces differently, and the proportion of time they spent looking at the eye region depended more on whether or not the faces were currently masked and not on whether or not the faces shown during familiarization were masked.

Finally, correlations between age in days and the proportion of time infants spend looking at the upper halves of the faces during familiarization in the four trial conditions are shown in Figure 6 (see Table S10 in the supplemental materials for full details). In general, when infants viewed unmasked faces (Figure 6, right column), the proportion of looking devoted to the upper halves of faces decreased with age, consistent with other findings of age-related changes in infants' bias to look at the eyes of unmasked faces. When viewing masked faces, however, the proportion of infants' looking to the upper halves of faces did not vary with age (Figure 6, left column), demonstrating that when the mouth is obstructed by a face mask even older infants do not devote much attention to the lower half of faces. Although there are age-related differences in infants' interest in the eye region (i.e., the upper half) of faces, infants' preference for the upper half of the face during familiarization was not related to memory performance (excluding extreme scores) in this sample (see Table S11 in the supplemental results for details), all $ps \geq .26$. Possible explanations for this pattern will be explored in the Discussion.

Discussion

We examined the effects of face masks on face memory in a sample of Western, North American infants born during the COVID-19 pandemic and tested during the first Omicron variant wave in the United States. In general, we observed that the presence of face masks did contribute to whether or not 6- to 9-month-old infants showed recognition memory for a face in the VPC task. Specifically, although infants showed memory when the test faces were unmasked, they did not show clear memory when the test faces were masked. Interestingly, face masks during the familiarization phase did not seem to be related to infants' performance in this task.

We also found that the presence of face masks influenced where infants looked when viewing faces in general. Although infants showed a bias to look at the upper regions of faces, the bias was stronger for masked faces than for unmasked faces. This is not surprising given that there is little to see in the lower halves of faces wearing nondescript black masks; however it does show that infants adapted their scanning patterns depending on the information available to them. In addition, as has been observed in other samples of Western infants (Oakes and Ellis 2013; Ayneto and Sebastian-Galles 2017), this bias to look at the upper regions of faces of unmasked faces decreased with age. At first glance, our findings might be taken to suggest that face masks present a challenge to infants, and that infants have difficulty remembering and distinguishing faces with face masks. Specifically, the 6- to 9-month-old infants in our sample showed less robust memory if the *test* faces were masked. However, aspects of our data show that infants do recognize and remember faces even when those faces are wearing face masks. Not only did infants in our sample show a significant novelty preference in the unmasked-familiarization/unmasked-test condition,

which is the "standard" condition that has been used in many studies, they also showed a significant novelty preference in the masked-familiarization/unmasked-test condition. Thus in our sample infants learned and formed a memory of a *masked* or *unmasked* face during the 12-s familiarization phase, and differentiated that face from a novel face when both the faces were *unmasked*. Thus, infants' memory performance does not appear to be hindered by the presence of a face mask during the learning (familiarization) phase, as long as the faces are unmasked during test. Infants' success on the masked familiarization/unmasked test condition also show that infants can encode a masked face and then generalize from one view of a face during familiarization to a different view of that face during test, as has been previously observed (e.g., Kelly et al. 2007, 2009; Otsuka et al. 2009).

Despite the fact that infants generalized from a masked face during familiarization to an unmasked face during test, they apparently did not generalize from the unmasked face during familiarization to the masked face during test. Why did we observe this asymmetry? One possibility is that our test phases were simply too short for infants to process the masked faces. That is, a 12-s familiarization presentation may have been sufficient for infants to process a single masked face (i.e., two copies of identical images), but two 5-s test presentations may have been too short to allow infants to sufficiently process two masked faces to recognize them as familiar or novel.

Alternatively, face masks may play a different role in infants' encoding of a new face (i.e., during familiarization) and their recognition that they have previously seen a face (i.e., during test). That is, the visual paired comparison procedure does not tap a single process, but rather requires that infants engage in multiple memory processes (Rose, Feldman, and Jankowski 2004). To successfully demonstrate memory, infants must first encode the face during the initial familiarization phase, recall that face during the test phase, and detect that the previously seen face was familiar. The presence of face masks may influence some of those processes more than others. Recall that responding in the masked-familiarization/ masked-test condition was ambiguous, and was related to infant age. Thus, in our sample it appears that masks may have influenced multiple memory processes. Future research may provide additional understanding into how face masks influence different memory processes and how these might change across development.

Finally, it is possible that these differences reflect infants' matching face parts to whole faces. Specifically, infants may be able to recognize a whole face (i.e., the unmasked face) when the initial input reflected only part of the face (i.e., the masked face), but unable to recognize a part of the face when the initial input reflected the whole face. This pattern is consistent with the "part-whole" effect (J. W. Tanaka and Farah 1993) and suggests holistic face processing. Holistic face processing is indicated by difficulty making judgments about facial regions (e.g., the top half) when that region is viewed in the context of a whole face, versus when it is viewed either misaligned from the rest of the face (the composite face effect) (Mondloch et al. 2007) or in isolation (James W. Tanaka and Gordon 2011). Young children have more difficulty matching and comparing the top halves of two faces when the top and bottom faces are aligned versus when they are not (i.e., the bottom half is shifted horizontally with respect to the top half) (de Heering, Houthuys, and Rossion 2007). In addition, children have difficulty recognizing isolated parts of faces they have

seen previously (James W. Tanaka et al. 1998). Thus, preschoolers and young children have difficulty recognizing isolated parts when they have previously seen those parts in a whole face, similar to how our infants failed to show a novelty preference to masked faces following familiarization with an unmasked face. Indeed, our 6 to 9 month range is the period when research using other methods suggests that infants begin to process faces holistically (e.g., Schwarzer, Zauner, and Jovanovic 2007; Scott and Nelson 2006). Future studies should more directly examine whether the effect of masks on infants' face processing reflects an interference with holistic processing.

It is important to point out that our conclusions are based on whether infants' novelty preference score was different from chance in each condition–our primary, preregistered analyses. However, none of our exploratory analyses revealed significant differences in novelty preferences *between* trial conditions. As is clear in Figure 3, infants' responding was highly variable in this task, as is typical for tasks like that used here (Rose, Feldman, and Jankowski 2004; DeBolt, Rhemtulla, and Oakes 2020), and it is possible that these differences would be uncovered in an experiment with more statistical power.

Recall that regardless of whether or not the face was masked, infants showed a bias to look at the upper or eye region of faces, perhaps reflecting a strategy that is advantageous for face processing. Several studies suggest that infants can use eyes to discriminate faces. Using the bubbles technique, Humphreys et al. (2006) found that 7-month-old infants could discriminate between a novel and familiar (mother's) face using the eyes. Vanderwert et al. (2015) found in a sample of 7-month-old infants from the Boston area that face specific ERP responses were associated with greater looking to the eyes during emotion processing. Moreover, infants' preference for eyes shifts with ages (Oakes and Ellis 2013; Ayneto and Sebastian-Galles 2017). In our sample, as has been observed in samples from similar cultural contexts, there was a decrease with age in infants' bias to look at the upper region of unmasked faces. Interestingly, this shift in bias did not translate to differences in infants' memory for faces. Specifically, although infants' preference for the top half decreased with age in the unmasked-familiarization/unmasked-test condition, in that condition there was no effect of age on infants' novelty preference and infants' upper half bias was unrelated to novelty preference. In contrast to other findings that the amount infants looked to the eyes is related to their learning of faces (Amso et al. 2010; Bolhuis, Kolling, and Knopf 2015), our data suggests that infants learned equally well regardless of how much they looked at the upper half of the faces.

However, we did find that infants' novelty preference was related to age when the face was masked during both familiarization and test, although infants' preference for the top half of the face did not vary with age in this condition. Thus, older infants seem to be better at remembering the face from their looking at the top half when the faces were masked during familiarization and test.

This pattern is consistent with the possibility that younger and older infants may form memories of faces using different information. That is, although the novelty preference score reflects memory for faces, it does not allow us to know whether that memory is based on learning about the same regions or using the same processes. For example,

Vogel et al. (2012) observed that although 5- and 9-month-old infants showed the same behavioral response to familiar-race faces (i.e., novelty preference), the behavioral response was associated with different underlying neural activity at the two ages. In the context of changes in how infants process faces (i.e., perceptual narrowing), it has been suggested that memory and discrimination by younger infants are driven by bottom-up processes and those same memory and discrimination by older infants are driven by more top-down processes (Hadley et al. 2014; Markant and Scott 2018). Thus, the age-related effects we observed here may be related to differences in how infants process faces, and there may be different developmental trajectories for processing during familiarization, or encoding, and test, or recognition. The present study is just a first step in understanding whether and how infants' processing of masked and unmasked faces may change over development.

Interestingly, we did not find differences as a function of experience with face masks (see supplemental materials). It is possible that our measure of face mask experience was not sensitive enough to adequately capture the true amount of variation in our infants' experience seeing face masks. Alternatively, face mask adherence may have been so pervasive in the community in which the data was collected that the amount of true variation in infants' experience seeing face masks was low, and that this is reflected in the responses. It is important to note that when the data was collected, there was a state mask mandate in place requiring people to wear masks in public indoor settings, so it is likely that overall mask endorsement was high overall during the time data was collected.

Like all empirical work, our study has limitations that have implications for the generalizability of our findings (Simons, Shoda, and Lindsay 2017). There are several factors that need to be taken into consideration. First, our sample included North American infants from middle class families. Their cognitive development — including their face processing — is shaped by everyday experiences (Rogoff, Dahl, and Callanan 2018; Ellis et al. 2017). Our findings may not generalize to infants from very different cultural contexts.

Second, data collection occurred when US COVID-19 case counts were extremely high (Cuadros et al. 2022) and local face mask mandates were in place. It is unknown what effect changes in face mask prevalence will have on findings like these. Moreover, because our study was conducted in a culture that prior to the pandemic did not routinely engage in face coverings, adults who are wearing masks may interact differently with infants than do adults in cultures where face coverings are more normative. Indeed, attitudes about face masks and face coverings vary between different cultures (Lu, Jin, and English 2021) and these attitudes may translate to different behavior while wearing masks. To explore these questions, future research is needed in cultures and contexts where face coverings are more widely adopted, and in some cases, woven into the cultural fabric.

Third, our stimuli comprise static images of unfamiliar faces, devoid of information like language, emotion, and facial movement typical of faces in real life. Future research could test whether the same patterns of results obtained in the present work would generalize to face stimuli that incorporate such additional cues. Lastly, the infants in our study completed a single trial per condition, possibly contributing to some ambiguity in some of the results.

Future studies should include more trials per condition to increase statistical power and the reliability of the findings (Byers-Heinlein, Bergmann, and Savalei 2021).

In summary, these results are a first step towards understanding the impact of face masks on infants' learning of and memory for faces. In general, we observed differences in infants' memory for faces only as a function of whether or not faces were presented during test, or the recognition phase of the trial, were masked. We did observe that infants' attention to faces during familiarization was impacted by masks, in that infants showed increased attention to the eye region while viewing masked faces compared to unmasked faces. However, older infants looked to the eye region less than younger infants, consistent with the literature showing that infants prioritize different facial regions across development. These findings advance our understanding of the features that are important for infants' learning and recognition of masked and unmasked faces, and aspects of our data demonstrate that infants do recognize and remember faces even when those faces are wearing face masks in a sample of North American infants.

Acknowledgments

References

Amso Dima, Fitzgerald Megan, Davidow Juliet, Gilhooly Tara, and Tottenham Nim. 2010. "Visual Exploration Strategies and the Development of Infants' Facial Emotion Discrimination." Frontiers in Psychology 1 (NOV): 180. [PubMed: 21833241]

Ayneto Alba, and Sebastian-Galles Nuria. 2017. "The Influence of Bilingualism on the Preference for the Mouth Region of Dynamic Faces." Developmental Science 20 (1). 10.1111/desc.12446.

Bennetts Rachel J., Humphrey Poppy Johnson, Zielinska Paulina, and Bate Sarah. 2022. "Face Masks versus Sunglasses: Limited Effects of Time and Individual Differences in the Ability to Judge Facial Identity and Social Traits." Cognitive Research: Principles and Implications 7 (1): 18. [PubMed: 35171394]

Bolhuis Jantina, Kolling Thorsten, and Knopf Monika. 2015. "Looking in the Eyes to Discriminate: Linking Infants' Habituation Speed to Looking Behaviour Using Faces." International Journal of Behavioral Development 40: 243–52.

Byers-Heinlein Krista, Bergmann Christina, and Savalei Victoria. 2021. "Six Solutions for More Reliable Infant Research." Infant and Child Development, December. 10.1002/icd.2296.

Carbon Claus-Christian, and Serrano Martin. 2021. "The Impact of Face Masks on the Emotional Reading Abilities of Children-A Lesson From a Joint School-University Project." I-Perception 12 (4): 20416695211038265.

Carragher Daniel J., and Hancock Peter J. B. 2020. "Surgical Face Masks Impair Human Face Matching Performance for Familiar and Unfamiliar Faces." Cognitive Research: Principles and Implications 5 (1): 59. [PubMed: 33210257]

Colombo John, Mitchell D. Wayne, and Horowitz Frances D. 1988. "Infant Visual Attention in the Paired-Comparison Paradigm: Test-Retest and Attention-Performance Relations." Child Development 59 (5): 1198–1210. [PubMed: 3168636]

Cuadros Diego F., Moreno Claudia M., Musuka Godfrey, Miller F. Dewolfe, Coule Phillip, and MacKinnon Neil J. 2022. "Association Between Vaccination Coverage Disparity and the Dynamics of the COVID-19 Delta and Omicron Waves in the US." Frontiers of Medicine 9 (June): 898101.

DeBolt Michaela C., Rhemtulla Mijke, and Oakes Lisa M. 2020. "Robust Data and Power in Infant Research: A Case Study of the Effect of Number of Infants and Number of Trials in Visual Preference Procedures." Infancy: The Official Journal of the International Society on Infant Studies 25 (4): 393–419. [PubMed: 32744759]

DeMets DL, and Lan KK 1994. "Interim Analysis: The Alpha Spending Function Approach." Statistics in Medicine 13 (13–14): 1341–52; discussion 1353–56. [PubMed: 7973215]

Ellis Ann E., Xiao Naiqi G., Lee Kang, and Oakes Lisa M. 2017. "Scanning of Own- versus Other-Race Faces in Infants from Racially Diverse or Homogenous Communities." Developmental Psychobiology 59 (5): 613–27. [PubMed: 28577346]

Fagan Joseph F. 1972. "Infants' Recognition Memory for Faces." Journal of Experimental Child Psychology 14 (3): 453–76. [PubMed: 4676459]

Fagan Joseph F., and Krahe McGrath Susan. 1981. "Infant Recognition Memory and Later Intelligence." Intelligence. 10.1016/0160-2896(81)90002-7.

Gava Lucia, Valenza Eloisa, Turati Chiara, and de Schonen Scania. 2008. "Effect of Partial Occlusion on Newborns' Face Preference and Recognition." Developmental Science 11 (4): 563–74. [PubMed: 18576964]

Grenville Emily, and Dwyer Dominic M. 2022. "Face Masks Have Emotion-Dependent Dissociable Effects on Accuracy and Confidence in Identifying Facial Expressions of Emotion." Cognitive Research: Principles and Implications 7 (1): 15. [PubMed: 35157157]

Grundmann Felix, Epstude Kai, and Scheibe Susanne. 2021. "Face Masks Reduce Emotion-Recognition Accuracy and Perceived Closeness." PloS One 16 (4): e0249792.

Hadley Hillary, Rost Gwyneth C., Fava Eswen, and Scott Lisa S. 2014. "A Mechanistic Approach to Cross-Domain Perceptual Narrowing in the First Year of Life." Brain Sciences 4 (4): 613–34. [PubMed: 25521763]

Adélaïde de Heering, Houthuys Sarah, and Rossion Bruno. 2007. "Holistic Face Processing Is Mature at 4 Years of Age: Evidence from the Composite Face Effect." Journal of Experimental Child Psychology 96 (1): 57–70. [PubMed: 17007869]

Hessels Roy S., Kemner Chantal, van den Boomen Carlijn, and Hooge Ignace T. C. 2016. "The Area-of-Interest Problem in Eyetracking Research: A Noise-Robust Solution for Face and Sparse Stimuli." Behavior Research Methods 48 (4): 1694–1712. [PubMed: 26563395]

Humphreys Kate, Gosselin Frédéric, Schyns Philippe G., and Johnson Mark H. 2006. "Using 'Bubbles' with Babies: A New Technique for Investigating the Informational Basis of Infant Perception." Infant Behavior & Development 29 (3): 471–75. [PubMed: 17138299]

Kelly David J., Liu Shaoying, Lee Kang, Quinn Paul C., Pascalis Olivier, Slater Alan M., and Ge Liezhong. 2009. "Development of the Other-Race Effect during Infancy: Evidence toward Universality?" Journal of Experimental Child Psychology 104 (1): 105–14. [PubMed: 19269649]

Kelly David J., Quinn Paul C., Slater Alan M., Lee Kang, Ge Liezhong, and Pascalis Olivier. 2007. "The Other-Race Effect Develops during Infancy: Evidence of Perceptual Narrowing." Psychological Science 18 (12): 1084–89. [PubMed: 18031416]

Lewkowicz David J., and Hansen-Tift Amy M. 2012. "Infants Deploy Selective Attention to the Mouth of a Talking Face When Learning Speech." Proceedings of the National Academy of Sciences 109 (5): 1431–36.

Lu Jackson G., Jin Peter, and English Alexander S. 2021. "Collectivism Predicts Mask Use during COVID-19." Proceedings of the National Academy of Sciences of the United States of America 118 (23). 10.1073/pnas.2021793118.

Luke Steven G. 2017. "Evaluating Significance in Linear Mixed-Effects Models in R." Behavior Research Methods 49 (4): 1494–1502. [PubMed: 27620283]

Markant Julie, and Scott Lisa S. 2018. "Attention and Perceptual Learning Interact in the Development of the Other-Race Effect." Current Directions in Psychological Science 27 (3): 163–69.

Mondloch Catherine J., Pathman Thanujeni, Maurer Daphne, Le Grand Richard, and de Schonen Scania. 2007. "The Composite Face Effect in Six-Year-Old Children: Evidence of Adult-like Holistic Face Processing." Visual Cognition 15 (5): 564–77.

Noyes Eilidh, Davis Josh P., Petrov Nikolay, Gray Katie L. H., and Ritchie Kay L. 2021. "The Effect of Face Masks and Sunglasses on Identity and Expression Recognition with Super-Recognizers and Typical Observers." Royal Society Open Science 8 (3): 201169.

Noyes Eilidh, and Jenkins Rob. 2019. "Deliberate Disguise in Face Identification." Journal of Experimental Psychology. Applied 25 (2): 280–90. [PubMed: 30730157]

Oakes Lisa M., and Ellis Ann E. 2013. "An Eye-Tracking Investigation of Developmental Changes in Infants' Exploration of Upright and Inverted Human Faces." Infancy: The Official Journal of the International Society on Infant Studies 18 (1): 134–48. [PubMed: 23525142]

Otsuka Yumiko, Konishi Yukuo, Kanazawa So, Yamaguchi Masami K., Abdi Hervé, and O'Toole Alice J. 2009. "Recognition of Moving and Static Faces by Young Infants." Child Development 80 (4): 1259–71. [PubMed: 19630907]

Pascalis Olivier, De Martin de Viviés Xavier, Anzures Gizelle, Quinn Paul C., Slater Alan M., Tanaka James W., and Lee Kang. 2011. "Development of Face Processing." Wiley Interdisciplinary Reviews: Cognitive Science. John Wiley & Sons, Inc. 10.1002/wcs.146.

Pascalis Olivier, de Haan Micheille, and Nelson Charles A. 2002. "First Year of Life? Is Face Processing Species-Specific During the." Science 296: 1321–23. [PubMed: 12016317]

Pascalis Olivier, de Haan Michelle, Nelson Charles A., and de Schonen Scania. 1998. "Long-Term Recognition Memory for Faces Assessed by Visual Paired Comparison in 3- and 6-Month-Old Infants." Journal of Experimental Psychology. Learning, Memory, and Cognition 24 (1): 249–60. [PubMed: 9438961]

Pascalis Olivier, Scott Lisa S., Kelly David J., Shannon RW, Nicholson E, Coleman M, and Nelson Charles A. 2005. "Plasticity of Face Processing in Infancy." Proceedings of the National Academy of Sciences of the United States of America 102 (14): 5297–5300. [PubMed: 15790676]

Pellicano Elizabeth, and Rhodes Gillian. 2003. "Holistic Processing of Faces in Preschool Children and Adults." Psychological Science 14 (6): 618–22. [PubMed: 14629695]

Pons Ferran, Bosch Laura, and Lewkowicz David J. 2015. "Bilingualism Modulates Infants' Selective Attention to the Mouth of a Talking Face." Psychological Science 26 (4): 0956797614568320–498.

R Core Team. 2019. "R: A Language and Environment for Statistical Computing." Vienna, Austria: R Foundation for Statistical Computing.

Rogoff Barbara, Dahl Audun, and Callanan Maureen. 2018. "The Importance of Understanding Children's Lived Experience." Developmental Review: DR 50 (December): 5–15.

Rose Susan A., and Feldman Judith F. 1987. "Infant Visual Attention: Stability of Individual Differences from 6 to 8 Months." Developmental Psychology 23 (4): 490–98.

Rose Susan A., Feldman Judith F., and Jankowski Jeffery J. 2004. "Infant Visual Recognition Memory." Developmental Review: DR 24 (1): 74–100.

Ruba Ashley L., and Pollak Seth D. 2020. "Children's Emotion Inferences from Masked Faces: Implications for Social Interactions during COVID-19." PloS One 15 (12): e0243708.

Schott Esther, Rhemtulla Mijke, and Byers-Heinlein Krista. 2019. "Should I Test More Babies? Solutions for Transparent Data Peeking." Infant Behavior & Development 54 (February): 166–76. [PubMed: 30470414]

Schwarzer Gudrun, Zauner Nicola, and Jovanovic Bianca. 2007. "Evidence of a Shift from Featural to Configural Face Processing in Infancy." Developmental Science 10 (4): 452–63. [PubMed: 17552935]

Scott Lisa S., and Nelson Charles A. 2006. "Featural and Configural Face Processing in Adults and Infants: A Behavioral and Electrophysiological Investigation." Perception 35 (8): 1107–28. [PubMed: 17076069]

Simons Daniel J., Shoda Yuichi, and Lindsay D. Stephen. 2017. "Constraints on Generality (COG): A Proposed Addition to All Empirical Papers." Perspectives on Psychological Science: A Journal of the Association for Psychological Science 12 (6): 1123–28. [PubMed: 28853993]

Stajduhar Andreja, Ganel Tzvi, Avidan Galia, Rosenbaum R. Shayna, and Freud Erez. 2022. "Face Masks Disrupt Holistic Processing and Face Perception in School-Age Children." Cognitive Research: Principles and Implications 7 (1): 9. [PubMed: 35128574]

Tanaka James W., and Gordon Iris. 2011. "Features, Configuration, and Holistic Face Processing." In Oxford Handbook of Face Perception, edited by Rhodes Gillian, Calder Andy, Johnson Mark, and Haxby James V. Oxford University Press.

Tanaka James W., Kay Joshua B., Grinnell Eliza, Stansfield Brent, and Szechter Lisa. 1998. "Face Recognition in Young Children: When the Whole Is Greater than the Sum of Its Parts." Visual Cognition 5 (4): 479–96.

Tanaka JW, and Farah MJ 1993. "Parts and Wholes in Face Recognition." The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology 46 (2): 225–45. [PubMed: 8316637]

Turati Chiara, Elisa Di Giorgio Lara Bardi, and Simion Francesca. 2010. "Holistic Face Processing in Newborns, 3-Month-Old Infants, and Adults: Evidence from the Composite Face Effect." Child Development 81 (6): 1894–1905. [PubMed: 21077871]

Vanderwert Ross E., Westerlund Alissa J., Montoya Lina, Mccormick Sarah A., Miguel Helga O., and Nelson Charles A. 2015. "Looking to the Eyes Influences the Processing of Emotion on Face-Sensitive Event-Related Potentials in 7-Month-Old Infants." Developmental Neurobiology 75 (10): 1154–63. [PubMed: 24962465]

Vogel Margaret, Monesson Alexandra, and Scott Lisa S. 2012. "Building Biases in Infancy: The Influence of Race on Face and Voice Emotion Matching." Developmental Science 15 (3): 359–72. [PubMed: 22490176]

Wass Sam V., Forssman Linda, and Leppänen Jukka M. 2014. "Robustness and Precision: How Data Quality May Influence Key Dependent Variables in Infant Eye-Tracker Analyses." Infancy: The Official Journal of the International Society on Infant Studies 19 (5). 10.1111/infa.12055.

Xiao Naiqi G., Quinn Paul C., Liu Shaoying, Ge Liezhong, Pascalis Olivier, and Lee Kang. 2015. "Eye Tracking Reveals a Crucial Role for Facial Motion in Recognition of Faces by Infants." Developmental Psychology 51 (6): 744–57. [PubMed: 26010387]
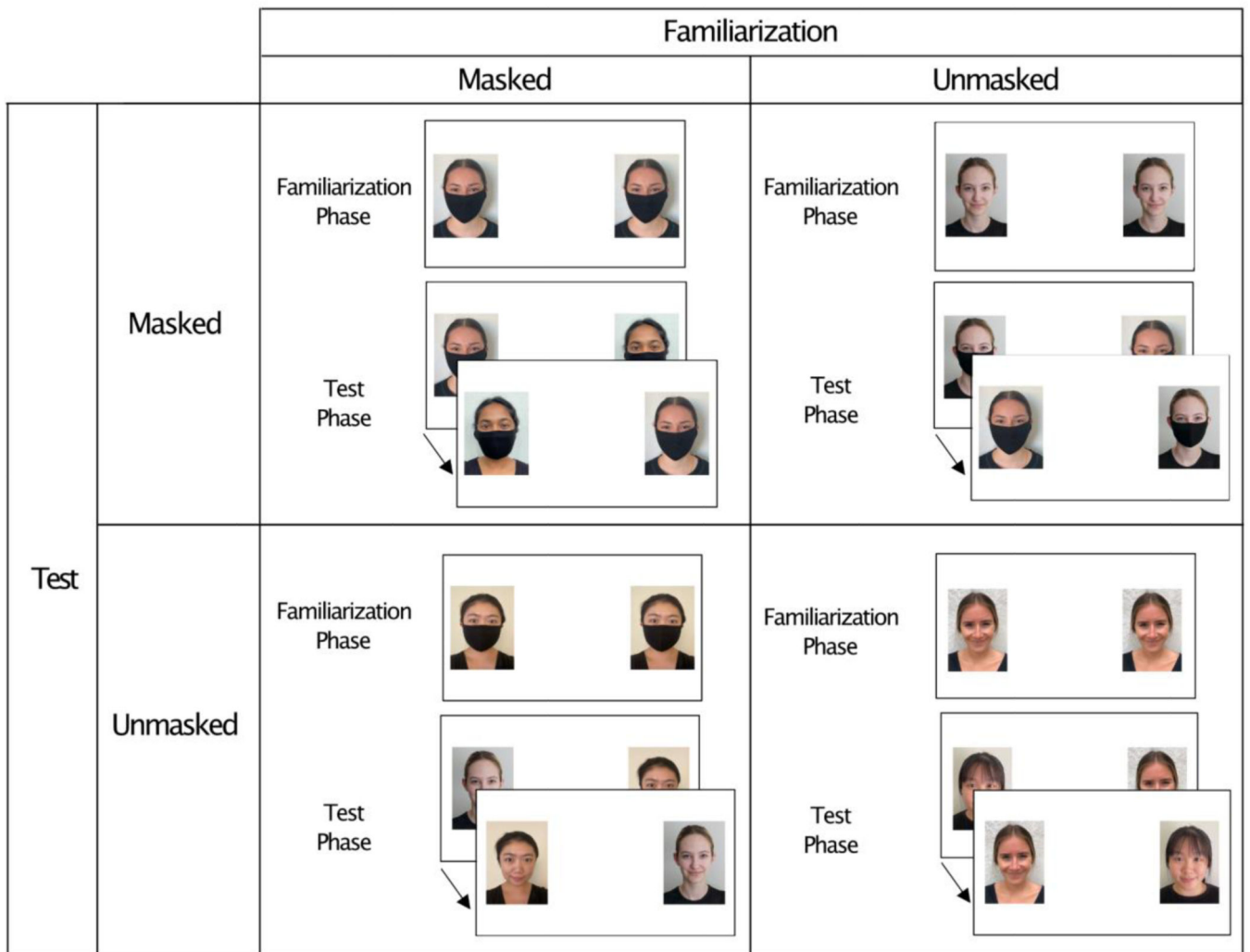
Figure 1.
Stimulus images and design. Each cell of the design depicts the familiarization stimulus, in which two pictures of the same woman's face were presented side-by-side, and the test stimuli, in which the previously familiarized face was paired with a novel face. The familiarization phase was a single 12-s presentation of the familiarization stimulus and the test phase involved 2 5-s presentations of the test stimuli (the left-right position of the stimuli reversed on the second presentation).

Figure 2.
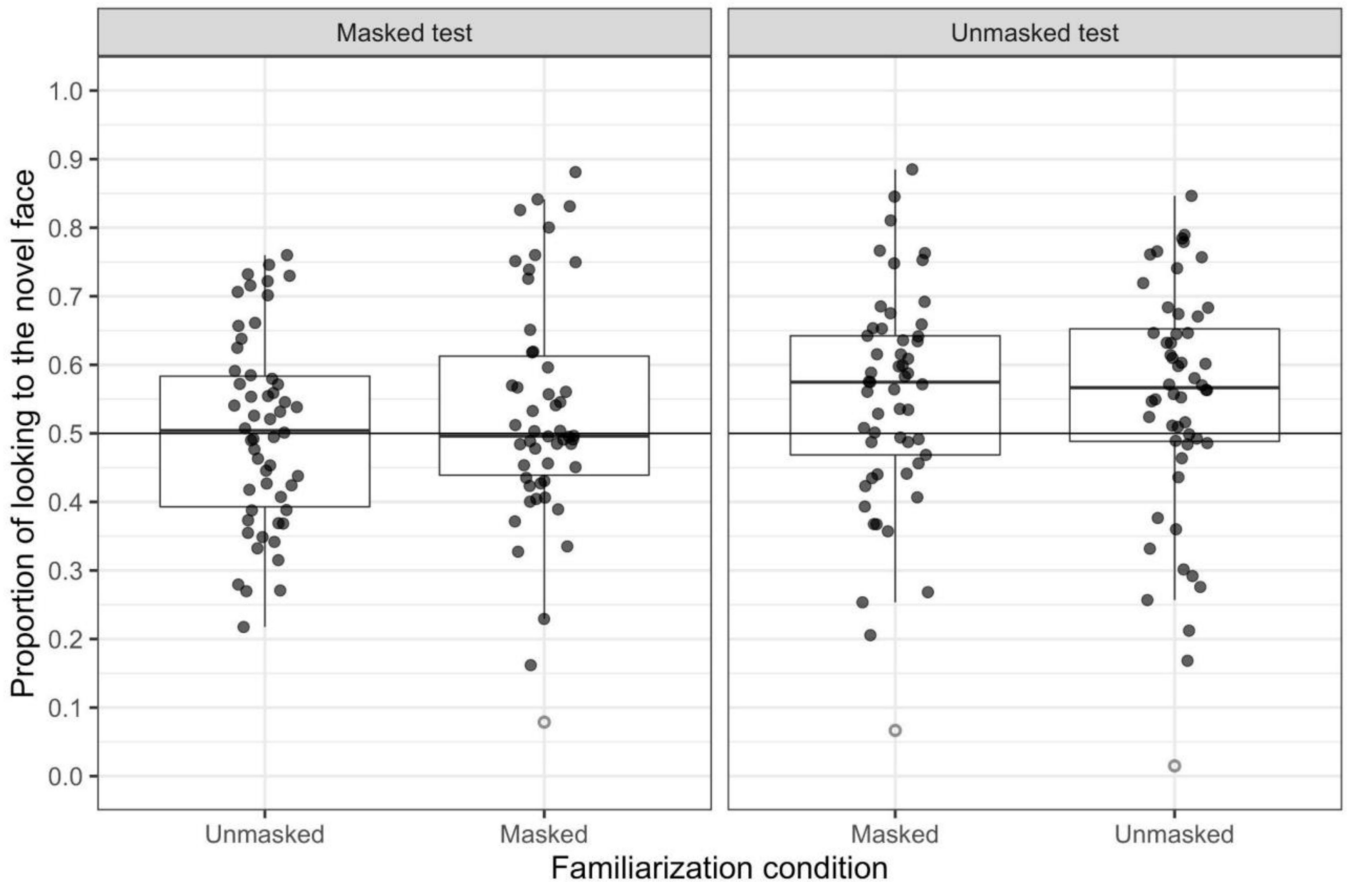Examples of face stimuli with upper and lower AOIs imposed.

Figure 3.
Trial-level data plots for each mask familiarization condition (x-axis) and mask test condition (labeled at the top of each panel). Each black dot represents a single participant's proportion of looking to the novel face on a single trial. The line bisecting the boxplots denotes the median novelty preference score; the lines extending from the boxplots extend no further than 1.5 X the IQR. The horizontal black lines denote chance responding (.50). The open gray circles indicate the extreme data points determined by the *boxplot* function in R across trial conditions (R Core Team 2019).
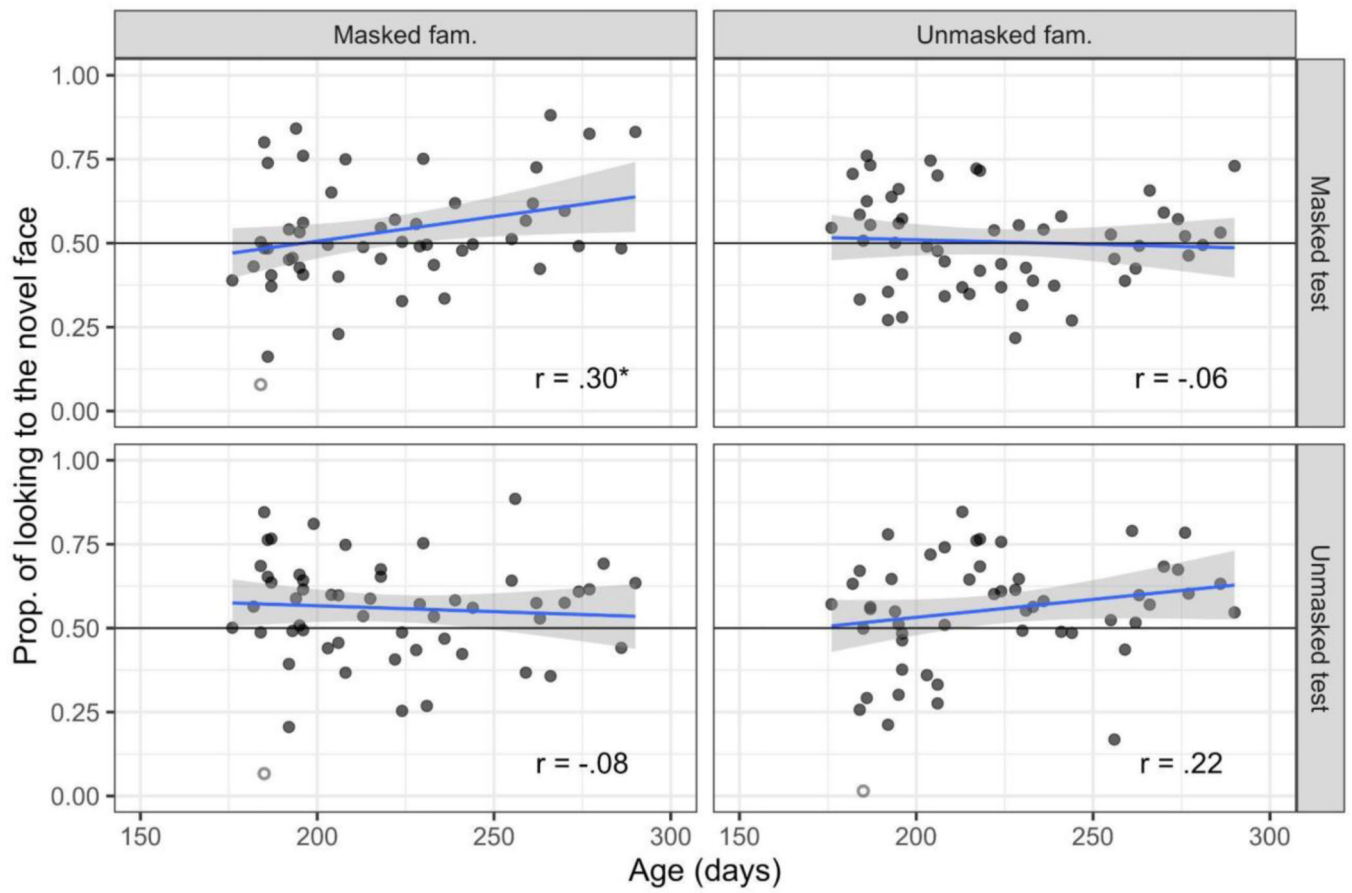
Figure 4.
Scatterplots of the correlation between novelty preference during test (y-axis) and age in days (x-axis), by condition. Extreme scores are indicated by the open circles and were not included in the correlation values listed in the plot. Correlations that were statistically significant, $p<.05$, are indicated by an asterisk.
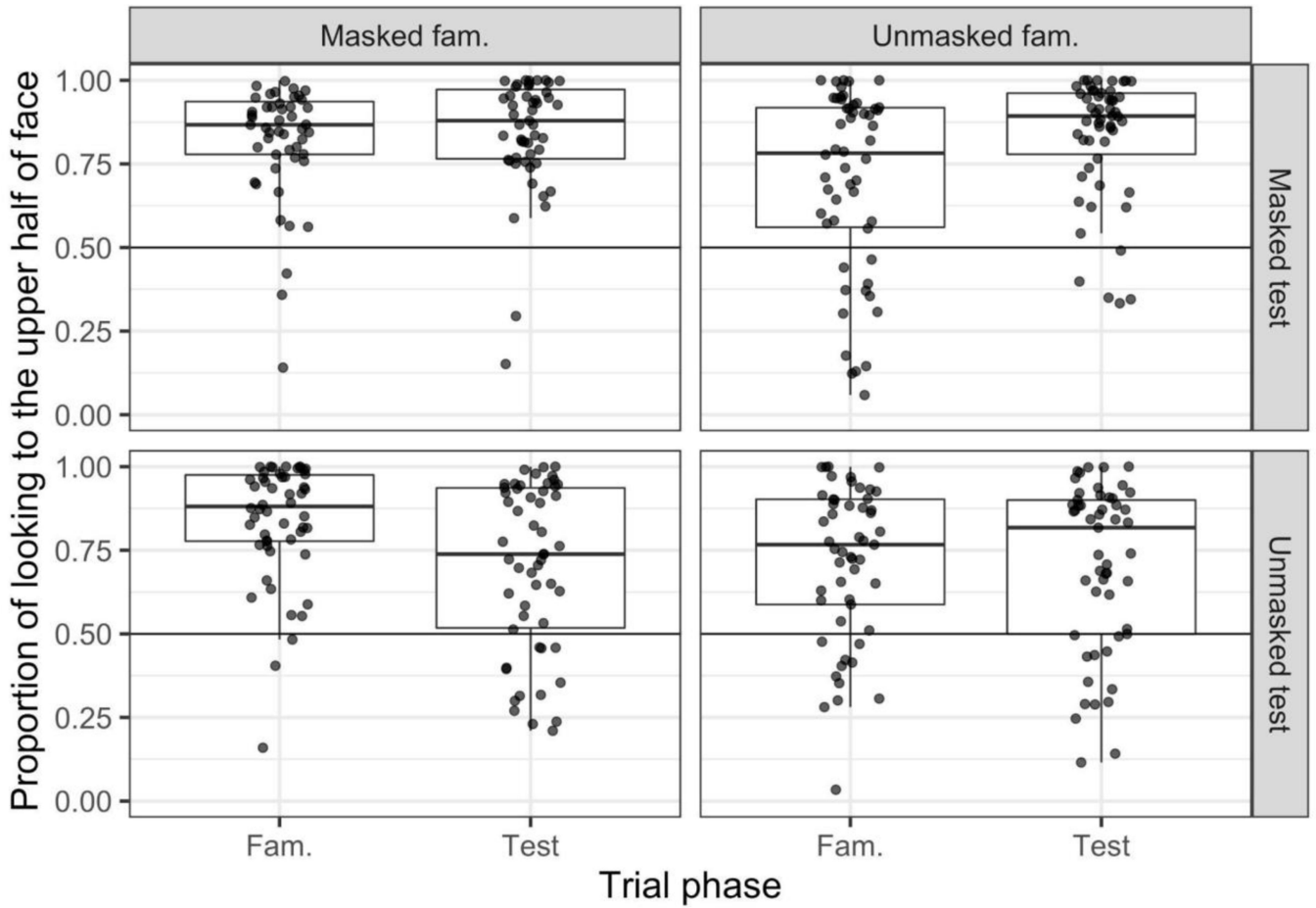
Figure 5.

The proportion of looking to the upper half of the face(s) during the familiarization and test phases (*x*-axis). In each plot, the individual black points represent a single infant's proportion of looking to the upper half of the faces. Note that for the familiarization phase, the proportion score reflects infants' looking to the upper half of the two identical faces whereas during the test phase the proportion score reflects infants' looking to the upper halves of the familiar and the novel face upper halves combined. The line bisecting the boxplots denotes the median novelty preference score; the lines extending from the boxplots extend no further than 1.5 X the IQR. The horizontal black lines denote chance responding (.50).
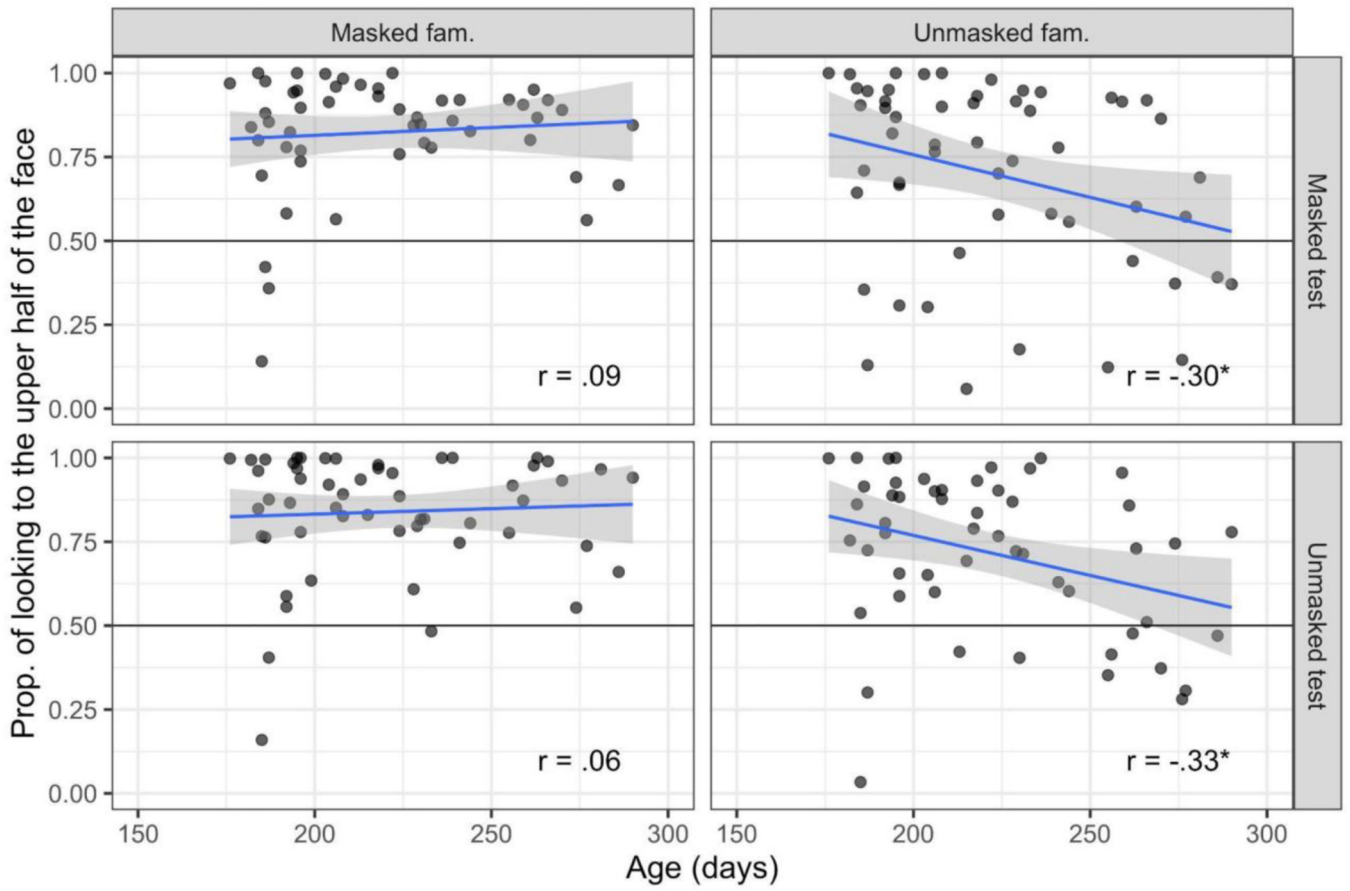
Figure 6.
Scatterplots of the correlation between preference for the upper half of the face during familiarization (y-axis) and age in days (x-axis), by condition. Correlations that were statistically significant, $p \le .03$, are indicated by an asterisk.

Table 1.

Proportion scores for the whole sample and for the sample excluding the three extreme scores.

Trial Condition	Including extreme scores					Excluding extreme scores				
	Mean (SD)	*df*	*t*	*p*-value	*d*	Mean (SD)	*df*	*t*	*p*-value	*d*
Unmasked-fam/Masked-test [*]	.50 (.14)	53	.21	.83	.03	.50 (.14)	53	.21	.83	.03
Masked-fam/Masked-test	.53 (.17)	50	1.12	.27	.16	.54 (.16)	49	1.59	.12	.22
Masked-fam/Unmasked-test	.55 (.16)	53	2.32	.03	.31	.56 (.15)	52	2.96	.005	.41
Unmasked-fam/Unmasked-test	.55 (.17)	52	1.93	.06	.26	.56 (.16)	51	2.56	.01	.35

Note. There were 3 trials (2 from one participant) identified as extreme scores. The number of trials included in the analysis are equal to the degrees of freedom + 1.

[*] No extreme scores were excluded from this condition.

Table 2.

Pearson correlation between novelty preference and age in days.

Trial Condition	Including extreme scores			Excluding extreme scores		
	df	*r*	*p*-value	*df*	*r*	*p*-value
Unmasked-fam/Masked-test	52	−.06	.66	52	−.06	.66
Masked-fam./Masked-test	49	.33	.02	48	.30	.04
Masked-fam./Unmasked-test	52	.00	.98	51	−.08	.59
Unmasked-fam./Unmasked-test	51	.26	.06	50	.22	.13

Note. No outliers were excluded in the unmasked-fam/masked test condition.