Developmental Psychology
2017, Vol. 53, No. 1, 38–49

0012-1649/17/$12.00 http://dx.doi.org/10.1037/dev0000230

Why are Faces Denser in the Visual Experiences of Younger Than Older Infants?

Swapnaa Jayaraman
Indiana University

Caitlin M. Fausey
University of Oregon

Linda B. Smith
Indiana University

Recent evidence from studies using head cameras suggests that the frequency of faces directly in front of infants *declines* over the first year and a half of life, a result that has implications for the development of and evolutionary constraints on face processing. Two experiments tested 2 opposing hypotheses about this observed age-related decline in the frequency of faces in infant views. By the people-input hypothesis, there are more faces in view for younger infants because people are more often physically in front of younger than older infants. This hypothesis predicts that not just faces but views of other body parts will decline with age. By the face-input hypothesis, the decline is strictly about faces, not people or other body parts in general. Two experiments, 1 using a time-sampling method (84 infants, 3 to 24 months in age) and the other analyses of head camera images (36 infants, 1 to 24 months) provide strong support for the face-input hypothesis. The results suggest developmental constraints on the environment that ensure faces are prevalent early in development.

Keywords: face, infant, development, environment, evolution

Human faces are central to human social interactions. Faces provide information about individual people's identity, about their age and gender, about their emotions, and about their intentions (Calder & Young, 2005; Cohen & Cashon, 2001; Haxby, Hoffman, & Gobbini, 2000; Kanwisher & Yovel, 2006; Zhao & Bentin, 2008). Further, human face processing has unique properties relative to other visual categories (e.g., McKone, Kanwisher, & Duchaine, 2007). Accordingly, theorists and empirical researchers have been interested in both the evolutionary constraints and developmental origins of face perception (Johnson & Morton, 1991; Kanwisher & Yovel, 2006; Scott, Pascalis, & Nelson, 2007). The result is an expansive literature that shows a protracted experience-dependent course of development (Scherf & Scott, 2012). In brief, shortly after birth, infants prefer to look at face-like over non-face-like stimuli (Fantz, 1963; Johnson & Morton, 1991; Cassia, Turati, & Simion, 2004), a bias that has been interpreted in terms of an in-born visual template for faces (Morton & Johnson, 1991). Newborns also show very early effects of experience, preferring to look at their mother's face over other female faces within days of birth that is dose-dependent on visual exposure to their mother (Bushnell, 2003). Over the first year, infant preferences to look at and abilities to discriminate faces change systematically, becoming increasingly tuned to the specific faces in their environment (see Pascalis et al., 2014; Scott et al., 2007). All in all, the broad developmental outline is like that seen more generally in many species-important abilities: strong constraints on outcome guided by biased internal processes that interact with an expected set of experiences (Aslin, 1981; Gottlieb, 1991).

The expected set of experiences for the development of human face processing is human faces. Although we know a great deal about the development of face processing in infancy, we know very little about face experiences. Are these experiences also constrained in ways that support visual specialization of face processing, and if so how? Without empirical data, theorists have often assumed—given their social importance—that faces are prevalent in human visual experiences throughout life (e.g., Calder & Young, 2005; Cohen & Cashon, 2001; Cohen Kadosh & Johnson, 2007; Haxby et al., 2000; Kanwisher & Yovel, 2006; McKone et al., 2007; Nelson, 2003; Tarr & Gauthier, 2000; Zhao & Bentin, 2008). However, in other domains, we know that the visual information available to infants changes in consequential ways with development itself (Bertenthal & Campos, 1990; Frank, Simmons, Yurovsky, & Pusiol, 2013; Gilmore, Raudies, & Jayaraman, 2015; Kretch, Franchak, & Adolph, 2014). Recent findings from studies using a new technology suggest that this may be the case for face experiences as well. The new technology uses mini head cameras to capture wearer-perspective scenes (Fathi, Ren, & Rehg, 2011; Pirsiavash & Ramanan, 2012; Schmitow & Stenberg, 2015; Smith, Yu, Yoshida, & Fausey, 2014). The special contribution of head cameras for developmental psychology, as noted by Braddick and

Swapnaa Jayaraman, Department of Psychological and Brain Sciences, Indiana University; Caitlin M. Fausey, Department of Psychology, University of Oregon; Linda B. Smith, Department of Psychological and Brain Sciences, Indiana University.

Correspondence concerning this article should be addressed to Swapnaa Jayaraman, Department of Psychological and Brain Sciences, Indiana University, 1101 East 10th street, Bloomington, IN 47405. E-mail: swapnaa@indiana.edu

Atkinson (2011), is that head cameras provide a means to capture developmentally indexed visual scenes.

The head camera studies suggesting developmental changes in the frequencies of faces in infant environments were designed for different purposes, conducted in different settings and geometries, and over different age ranges of infants. However, in aggregate they indicate a clear developmental pattern. Findings from studies using head cameras at-home (Fausey, Jayaraman, & Smith, 2016; Jayaraman, Fausey, & Smith, 2015; Sugden, Mohamed-Ali, & Moulson, 2014) as well as in-laboratory settings (Frank et al., 2013; Pereira, Smith, & Yu, 2014; Yu & Smith, 2012) show that faces are much more frequent in the captured images for infants younger than 4 months of age than they are for toddlers. For all infants, the faces in the images appear to be most typically frontal (or near frontal) views within 4 feet of the infant (Jayaraman et al., 2015). However, the *frequency* of faces in the head camera views decreases systematically and markedly over the first year and a half of life. What are the origins of this developmental pattern in the frequency of faces in infant-perspective scenes? The present study was designed to test two hypotheses about why the frequency of faces in infant-perspective views declines with age.

The first hypothesis, proposed by Jayaraman, Fausey, and Smith (2015), is that people are more in view for young babies because very young babies need continuous care for feeding, for cleaning, for temperature and emotional regulation—all activities that seem likely to put people in front of infants. In contrast, older infants or toddlers are not so needy and thus caretakers need not be so often directly in front of the child. The hypothesis is not that people are more often spatially near younger than older infants, but that people are more often oriented with respect to the infant in ways that put them in front and thus in view of the infant. This distinction between being near and being in-view is illustrated in Figure 1. All individuals are physically near the infant, but the person sitting on the side of the infant is not in the infant's view. Although there is no empirical evidence on the matter, we suspect (and test in Experiment 1) that people are physically near older as well as younger infants. Older infants and toddlers spend their waking time actively exploring their environments and the evidence suggests that they do so without consideration of the dangers (Franchak & Adolph, 2012). Thus, older infants need constant monitoring lest they stick their fingers in sockets or topple down stairs. However, young-infant caretaking and older-infant monitoring may put the parent in different spatial relations with respect to the infant's head and eye. In sum, by what we call the people-input hypothesis, the greater frequency of faces in the view of younger than older infants is a byproduct of young infants' immaturity and need for caretaking, conditions which place people (and their faces) frequently in front of the infant. One implication of this hypothesis is that the greater frequency of faces in the visual experiences of young infants derives from a *non-face-specific* constraint on where people are relative to the infant. This hypothesized environmental constraint might be expected to interact with face-specific biases that sustain looks to face-like configurations (Morton & Johnson, 1991). But by the people-input hypothesis, the developmental constraint that makes faces more prevalent in the input for younger than older infants is not face specific but is about the frequency of people in the infant's visual field.

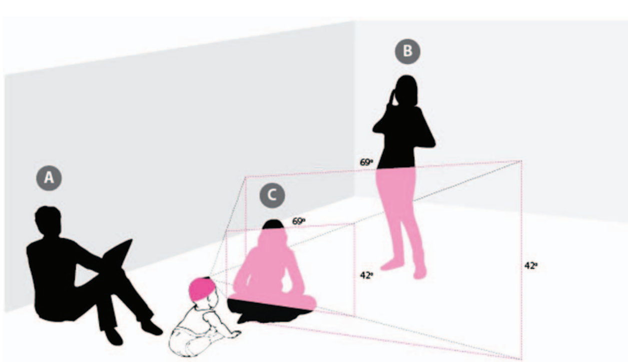


Figure 1. Three measures of people present around the infant. Time sampling of people in the same room as infant (captures Persons A, B, and C), head camera recording of people in infant's view (captures Persons B and C, loses person A), and head camera recording of *faces* in infant's view (captures the face of Person C only). See the online article for the color version of this figure.

The second hypothesis, the face-input hypothesis, proposes that there are developmental constraints on visual environments that specifically increase the frequency of faces for very young infants. As Gottlieb (1991) argued in his theory of environmental canalization, evolution may tightly constrain developmental outcomes not through internal mechanisms in the developing organism but by constraining the environment and the timing of specific experiences in that environment. From this perspective, the greater density of faces in the visual fields of younger than older infants could reflect a strong constraint on specifically *early* face experiences, a result that would suggest that the density of early face experiences is particularly critical to the development of face processing. The key prediction of the face-input hypothesis is that the frequency of people in the infant-perspective scenes does not decline with age, only the frequency of faces. This prediction is consistent with prior head camera studies of toddler toy play in the laboratory which have consistently reported a lack of faces in the toddler head camera (Yoshida & Smith, 2008; Yu & Smith, 2013) despite the fact that the parents of the toddler are in the room and actively engaged with the child. The prediction also fits with head camera (Fausey et al., 2016; Yoshida & Smith, 2008; Yu & Smith, 2013) as well as infant looking studies (Frank et al., 2013) that suggest an increased attention to and interest in hands (and hand actions) by older infants. However, because these studies measured brief durations of behavior in the laboratory in constrained visual contexts (Yoshida & Smith, 2008; Yu & Smith, 2013) or focused on hands rather than people and all their body parts more generally (Fausey et al., 2016), they do not provide the relevant information to test the people- versus the face-input hypotheses. In sum, by the face-input hypothesis, in contrast to the people-input hypothesis, the age-related decline in the frequency of faces in infant visual environments neither co-occurs with a decline in people near the infant nor with a decline in the people in view of the infant. Rather, by the face-input hypothesis, the decline is strictly about how frequently faces are positioned in front of infants.

These two contrasting hypotheses matter for multiple questions relevant to understanding the input in the development of face processing. For example, if the people-input hypothesis were correct, it would suggest a more critical role for early intrinsic biases that encourage young infants to look to faces over other body parts in the development of face processing. The face-input hypothesis, if correct, would suggest the perhaps overdetermined density of face input through redundant environmental and intrinsic constraints. The two hypotheses also have different implications for variations in face experiences and their consequences as a function of limited early vision (e.g., Le Grand, Mondloch, Maurer, & Brent, 2001), institutionalization (Moulson, West3erlund, Fox, Zeanah, & Nelson, 2009), and possibly with culture. These are considered in the General Discussion.

To test the people-input versus face-input hypotheses, we employed three measures of the people near infants in their own home. The measures were designed to distinguish the three cases illustrated in Figure 1: Person A is in the same room as the infant but not directly in front of the infant and thus not in the infant's view. Person B is in the room, her body is at least partly in the infant's view but her face is not. Person C is in the room and spatially positioned so that her face is in view of the infant. In Experiment 1, we used a time-sampling approach to measure the people in the same room as the infant. This measure captures persons in all three spatial arrangements—A, B, and C—in Figure 1. If the faces in the visual environment decline with age because the people spatially near the infant decline, then the number of people near the infant by this measure should decline with age. This experiment thus provides critical information lacking in the literature and information needed to interpret the direct test of the people- versus face-input hypotheses in Experiment 2. For that test, we used mini head cameras to measure the frequency of people and the frequency of faces in the infant view. To measure the frequency of people in view, we counted a person in view if *any* body part—hands, knees, edge of shoulder, whole body, face—was in the infant head camera image. This measure captures persons B and C in the figure. Again, if the people-input hypothesis is correct and the observed age-related decline in the frequency of faces in infant-perspective views is due to the frequency of people positioned within the infant's view, then the frequency of any body part being in view should decline with age. Finally, we measured the frequency of faces in infant-perspective views, Person C in the figure, using similar procedures to those used in recent head camera studies of the faces in infant visual environments (Jayaraman et al., 2015; Kretch et al., 2014; Sugden et al., 2014; Yu & Smith, 2013). If the face-input hypothesis is correct, the reported age-related decline in faces in infant-perspective views is not about the people near the infant, not about the people in view, but is specifically about faces and thus *only* this measure should decline with age.

The two experiments use very different methods and thus different approaches to measuring age-related changes. Experiment 1 is a time-sampling study of the people in the same room over the course of a whole day for 84 infants in a cross-sectional study of five age groups, 3, 6, 12, 18, and 24 months. Experiment 2 analyzes head camera images from a corpus composed of images collected from 36 infants aged 1 to 24 months, with an average of 4 hr of head camera video from each infant. In Experiment 2, age is treated as a continuous variable given the large amount of data per individual participant.

Experiment 1

Time sampling is a method in which observation is intermittent: brief intervals of time are repeatedly sampled for the property of interest. Across many disciplines this has been shown to provide good estimations of distributional properties over extended time periods that cannot be measured in their entirety (Bakeman, 1997). Here we use the method to measure the number and frequency of people in the same vicinity as the infant.

Method

Participants. Parents of 84 infants (34 female) aged 3, 6, 12, 18, and 24 months (19, 14, 15, 18, and 18 infants in each group, respectively) volunteered to participate. Infants were within +/-2 weeks of their assigned group age. The sample of infants was broadly representative of Monroe County, Indiana: 84% European American, 5% African American, 5% Asian American, 2% Latino, 4% Other, and consisted of predominantly working- and middle-class families.

Procedure. For initial instructions, parents were either invited into the laboratory or were visited at their homes. They were given log sheets and were instructed by a researcher on how to use them. The sheets contained tables spanning a 24-hr period with rows marking a query at 30-min intervals. At each queried moment, parents were asked to indicate the state of affairs for that moment: How many people were in the same room (vehicle or outside play area) as the infant such that the infant could—if they turned their body—see (any view of) a person (that is the person was not behind a wall). Parents were instructed not to include themselves as one of the people in the room if their presence was due only to the task of answering the query. Parents were also asked to indicate at each time moment if the infant was awake. In case parents missed a query, they were told to record the momentary state when they first noticed the missed query (marking the time) and then proceed with the scheduled queries. In providing these instructions, the experimenter considered various scenarios, questions about layouts and special cases, and answered them according to a set of rules. These rules were in accordance with the principle that if the infant altered body posture but did not move (crawl or walk) to a new location in the space (e.g., with respect to Figure 1, turned their head), the person would be (at least in part) in view. Parents were also asked about the daily routines and child care of their infants. If there was a caretaker other than the parent, the caretaker was asked to fill out the survey. Some infants attended some group activity occasionally (e.g., a once-a-week toddler music class) or regularly (some form of small-group day care). For queries that overlapped with those time intervals, parents were asked to record the number of the people in the program (e.g., a small-group day care with five other children, a teacher, and an aide should be recorded as seven people in the vicinity), a procedure that potentially overestimates the people (at any moment) in the vicinity of the infant. Parents were encouraged to call a researcher if they had questions during observation and many used that option. Upon completing the time sampling, parents returned the log sheets and were interviewed by the experimenter about their entries. Because the number of people in the vicinity of the infants could vary as a function of week days and week ends, parents of infants at each age group were randomly assigned a day of the week (Monday–

Sunday). Families received a small gift (book or t-shirt) as appreciation for their participation.

Results

Log sheets from each of the infants contained potentially 48 entries, one for each half hour of the 24-hr recording period. The number of entries differed only as a function of infant waking hours as all parents completed all wake hour queries (and did so within a half hour of the queried time). Because the central question for this experiment is how and whether the number of people in the vicinity of infants changes with infant age, we used a one-way analysis of variance with the five age groups as the between-subjects factor for all analyses.

The mean number of hours infants spent sleeping decreased with age, $F(4, 80) = 4.05$, $p < .01$. The youngest infants slept for an average of 14 hr thus yielding fewer data points than the oldest infants who slept for 12 hr of the day. These findings are consistent with infant sleep patterns across the first 2 years of life (Wooding, Boyd, & Geddis, 1990). For the following analyses, we used proportions of awake time that people were present, the mean numbers of people present at any awake point in the day, as well as the total cumulative number of people (uncorrected for wake time) who were reported to be in the room at each query.

Figure 2 shows histograms of the proportion of Times 0, 1, 2, 3, or more people were reported to be in vicinity of the infant. Table 1 provides the mean, median, mode, total number, and range for each of the age-level distributions, as well as the proportion of times at least one person was the in the room with the infant. To assess whether these distributions of people around infants changed with age, we calculated the mean, median, and mode for each infant's distribution and submitted these to separate one-way analyses of variance with age group as the between-subjects factor. There were no reliable differences in these measures of central tendency as a function of age group: Mean, $F(4, 80) = 0.91$, $p = .47$; Median, $F(4, 80) = 0.354$, $p = .55$; Mode, $F(4, 80) = 2.5$, $p = .12$. There was also no age-related change in minimum number of persons around infants, $F(4, 80) = 1.26$, $p = .29$, nor maximum number of persons, $F(4, 80) = 1.25$, $p = .30$. However the *total* number of persons, which is directly related to the number of hours the infants are awake, increased reliably with age, $F(4, 80) = 6.76$, $p < .05$, a result in the opposite direction of the reported age-related decline in faces in infant-perspective head camera scenes. Finally, as is evident in Table 1, there was at least one person in the vicinity of the infants nearly all the time and this did not vary with age, $F(4, 80) = 0.72$, $p = .58$. In brief, there is no decline with age in the number or frequency with which people are reported to be in the vicinity of the infant. Thus, the number of people spatially near infants—in the same room—will not account for the age-related decrease in faces found in head camera studies.

The age-related *increase* in total people in the vicinity of the infant over the course of the day is likely related to longer waking hours and perhaps also to increased group activities for older than younger infants. At least once during their recorded hours, 48 infants (57%) in the sample were in some group setting where more than three other individuals were present (e.g., playgroup, toddler music class, at the park, in a restaurant, etc.). On average, these infants who had some group activity spent less than 8% of the queried times in such settings. When considered as a group, the 84 infants from the dataset spent only 4% of their queried time in large group settings. Nevertheless, the frequency of time spent in settings with more than three individuals present increased with age, $F(1, 82) = 4.62$, $p < .05$, a pattern consistent with national data on daycare practices in the United States (Hofferth, 1996; Kreader, Ferguson, & Lawrence, 2005; Morrissey & Banghart, 2007). However, the critical fact is this: There is no decrease in the frequency of people in the vicinity of children over the first 2 years of life and thus a decline in potentially see-able people (in the same room) cannot account for the age-related decline in faces in infants' immediate visual environments.

Experiment 2

As illustrated in Figure 1, the people spatially near the infant need not be in view of the infant. To be in-view, those people need to be in front of the infant's face. The question for Experiment 2

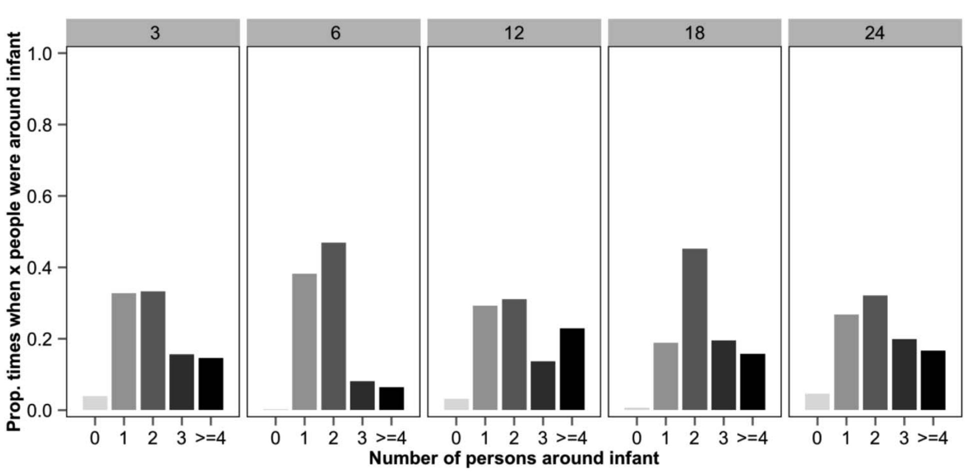


Figure 2. Histogram of persons present around infant. Proportions of times when 0, 1, 2, 3, or 4 or more persons were present around the infant are shown in separate panels for the infants in the five age groups. Panels are labeled (top bar) by the age in months of children in that group.

Table 1
Summary of Number of People Around Infants Recorded in Time Sampling Log Sheets

Age	Mean	Std. Dev.	Median	Mode	Range	Total	Prop. of times when at least one person was present
3 months	2.17	1.38	2.2	1.9	.68–4.10	44.0	.961
6 months	1.85	.95	1.9	1.8	1.00–3.14	39.2	.997
12 months	2.41	1.57	2.2	2.2	1.06–3.88	56.5	.970
18 months	2.37	1.13	2.3	2.1	1.11–4.00	56.0	.992
24 months	2.30	1.42	2.3	2.4	1.00–4.22	55.4	.954

Note. Each measure was calculated for each subject and then averaged for the subjects within an age group.

is whether the age-related decline in faces in the infant view is due a general decline in the frequency of people positioned in front of the infant or, more specifically, a decline in the frequency of faces located in front of the infant's own face. To answer this question, Experiment 2 analyzed over 100,000 images sampled from a corpus of over 16 million infant head camera images.

Because the data analyses in Experiment 2, and the previous findings of an age-related decline in the visual frequency of faces, involve analyzing head camera images, it is cogent to provide information on what is known about the limitations of this method. These are becoming increasingly better understood with the rising use of head cameras (Aslin, 2009; Frank et al., 2013; Kretch et al., 2014; Schmitow & Stenberg, 2015; Schmitow, Stenberg, Billard, & von Hoffsten, 2013a; Smith et al., 2014). One potential limit concerns the relation between eye and head direction, because head cameras measure head direction and not gaze direction. Further, the field of view (FOV) of the camera itself is less than the field of view of the infant and is particularly limited in the vertical and horizontal directions, so that, in principle, eye-gaze could be outside of the captured image. As discussed in a recent article on head cameras (Smith et al., 2014, see also Schmitow & Stenberg, 2015), there are two facts that mitigate against these limitations. First, in active viewing (not watching screens), infants as well as adults typically turn heads and eyes in the same direction to attend to a visual event (e.g., Ballard et al., 1992; Bambach, Crandall, & Yu, 2013; Bloch & Carchon, 1992; Daniel & Lee, 1990; Yoshida & Smith, 2008) and sustained visual attention is associated with aligned heads and eyes (Pereira et al., 2014; Ruff & Lawson, 1990). Although eyes often lead heads in directional shifts of visual attention (and although heads undershoot eyes given extreme changes in gaze direction), differences in head-eye direction are usually resolved in less than 500 ms in infants (Schmitow, Stenberg, Billard, & von Hofsten, 2013b; Yoshida & Smith, 2008). Head-mounted eye-tracking studies also show that aligned heads and eyes—that is, fixations to the center of the head-centered image—strongly characterize freely moving toddlers' viewing behavior (Bambach et al., 2013). The overwhelming predominance of gaze centered within the head camera image (see Bambach et al., 2013) reduces the likelihood of missed content due to momentary shifts in eye-gaze and the FOV of the camera itself. Thus, the assumption that underlies the use of the head camera to measure the content of infant visual experiences is that if the sample of head camera images is large enough, then the observed regularities in content may be assumed to characterize the content of scenes typically in front of both the heads and eyes of infants.

Method

Participants. Parents of 36 infants (19 female) between ages 1–24 months volunteered to participate in the head camera study. Participants were recruited from the same population and through the same methods as Experiment 1. Families received $30 gift card and a small gift (book or t-shirt) as appreciation for their participation. The head camera data generated by these infants is part of a growing corpus of infant-perspective scenes collected to understand the natural statistics of early visual experience and how it changes with age. Thus, 22 of the infants in the present analysis (all younger than 13 months of age) contributed head camera data to the earlier study (Jayaraman et al., 2015) that first reported an age-related decline in the faces in view within the first year; 34 of the infants contributed head camera images to the study (Fausey et al., 2016) that analyzed hands in infant-perspective scenes. Two infants in the present study were new to the corpus. We describe the sampling procedure from the corpus of head camera images below. We selected different unique samples of images from each of the 22 infants in the Jayaraman et al. (2015) study. Thus, although the evidence we report on the frequency of faces in these infant-perspective scenes is not a completely independent replication of the age-related decline in faces reported in the Jayaraman et al. (2015) study, the analyses provide additional support for that prior finding by adding 12 more infants and by sampling different frames for analysis from the overlapping infants. The strictly new contribution is the analysis of the people in these infant-perspective scenes and the comparison of age-related changes in the frequency of faces in the images to the frequency of people in the images.

Head camera. The principle motivation behind the choice of the specific head camera used to collect the corpus of images was not to influence—by our presence during recording—the recorded scenes. Therefore, our camera system was selected to be easy to use by parents without any technical help, to be light enough for infants to wear and forget about, to not heat up, and to be safe. The camera was mounted onto a snug hat and recorded the video on an internal storage system. The entire system weighs 2 oz (pack of gum). The camera has a 75° diagonal field of view (41° vertical FOV, 69° horizontal FOV), capturing a broad view of what is directly in front of the infant's head (see Figure 1). Parents were given two hat-camera systems, each with a recording capacity of 3–4 hr of video at the rate of 30 frames per second.

Instructions. Invitation to participate and debriefing procedures were structured as in Experiment 1. Parents were instructed in the use of the camera in the preliminary visit. They were told that we were interested in the visual experiences of infants, that all aspects of the infant's day were relevant, and that they should not alter their daily schedule for recording. They were asked to record during the infants' waking hours and to try to capture 4 to 6 hr of video during typical daily activities and contexts and were allowed multiple days to do so.

Coding. Videos from each participant were screened for privacy and for accidental recordings when the camera was not on the baby (1.5% of total recording) and the remaining videos were further processed to extract individual frames and create a set of images sampled once every 5 s from the videos (see Figure 3). The sampled images were organized into randomly ordered sets of 100 for presentation to trained coders naïve to the hypotheses under consideration. Coders were trained on a set of nine instruction images. Four coders answered a single yes/no question about each image. For the "person in view" measure, the coder was asked "Do you see a person (human body or any body part) in the picture?" A frame was deemed to contain a person if at least three out of four coders agreed. For 95.9% of the frames, at least three coders were in agreement that the frame did or did not contain a person. To ensure that the body parts were not the infant's own body, on a separate coding pass, coders were asked "Does this body part belong to the infant wearing the camera?" Images for which coders answered "yes" were further coded for "Do you see anybody in this image other than the infant's own body or body part?" Images that only contained body parts of the infant wearing the head camera with nobody else in view were removed from subsequent analyses (1.6% of the images). In a separate coding event, the entire sample was coded, using the same procedure, with the question being "Do you see a face or a face part in this image?" for which the agreement (at least three of coders said "yes" or at least three of four coders said "no") was 95.89%. Sampling at 1/5 Hz does not appear to have any age-related biases and appears to be sufficiently dense enough to capture major regularities. In particular, a coarser sampling of images at 1/10 Hz yielded the same age-related patterns; further, a sampling of a different set of images at 1/5 Hz but that used new starting points yielded the same patterns of age-related changes (see also Jayaraman et al., 2015). Figure 3 provides examples of frames sampled at 1/5 Hz and face-in-view and person-in-view coding.

We also report the distance of faces from the infants under 3 months of age (using the same image sampling and coding reported in Jayaraman et al., 2015). To estimate the distance of each face from the head camera, trained coders matched faces (and parts of faces) to size templates that were generated by determining head camera image size for an adult female face at 1-foot increments from the head camera. A second coder independently coded 20% of the sample, and coder agreement exceeded 98%.

Total recording hours from each infant ranged from 2.5 to 7 hr (with the exception of the youngest infant who provided only 15 min). The hours of recording obtained from each infant did not vary systematically with age, $R^2 = .023$, $F(1, 31) = 0.83$. $p = .37$. Across all infants, the average usable recording time was 4.2 hr. The total corpus, over 151 hr of video sampled at 1/5 Hz, consisted of 108,984 frames; the mean number of frames (the data points determining the frequency of faces and people in the view) contributed by individual infants was 3,027.

Results

The overall percentage of frames that contained a person other than the infant was 39.6%. That is, other people were in view, on average, for nearly 24 min per recorded hour. To analyze the age-related properties of people and faces in these views, we used linear regression. Figure 4A shows each individual infant's proportion of frames that contained any body part of another person as a function of age. Age and the frequency of a person in view were not related, $R^2 = 0.001$, $F(1, 34) = 0.004$, $p = .95$. People are equally likely to be in view across the age range, a result that contradicts the people-input hypothesis.

The percentage of times a face or part of a face (which is a subset of any body part queried in the previous question) was in the head camera image was 17.5%. Figure 4B shows this proportion declines as a function of the age of the infant. Consistent with previous analyses (Jayaraman et al., 2015), faces were not uniformly frequent across ages. Faces were available in view about 23.2% of the times for the youngest seven infants (3 months old or younger), but declined reliably with age, $R^2 = 0.29$, $F(1, 34) = 13.83$. $p < .001$, to about 8.83% for the oldest seven infants (over 18 months old).

Overall, the results provide clear evidence for the face-input hypothesis: The decline with age in the frequency of faces in head camera images appears specific to faces and not to the presence of people more generally. Figure 4C makes clear the dramatic dif-

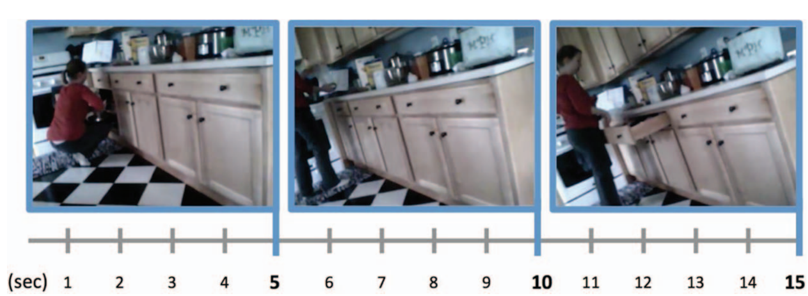


Figure 3. Downsampled frames. This figure illustrates frames sampled at the rate of 1 every 5 seconds (the sampling rate used for analysis) from the head camera video. Coders were asked to respond "yes" if any part of the body (no matter how small) or any part of the face (from any perspective) was in view. Thus these three frames would be coded as "yes-yes-yes" for person-in-view question and "yes-no-yes" for the face-in-view. The individual featured in these images has provided signed consent for her likeness to be published by the authors.

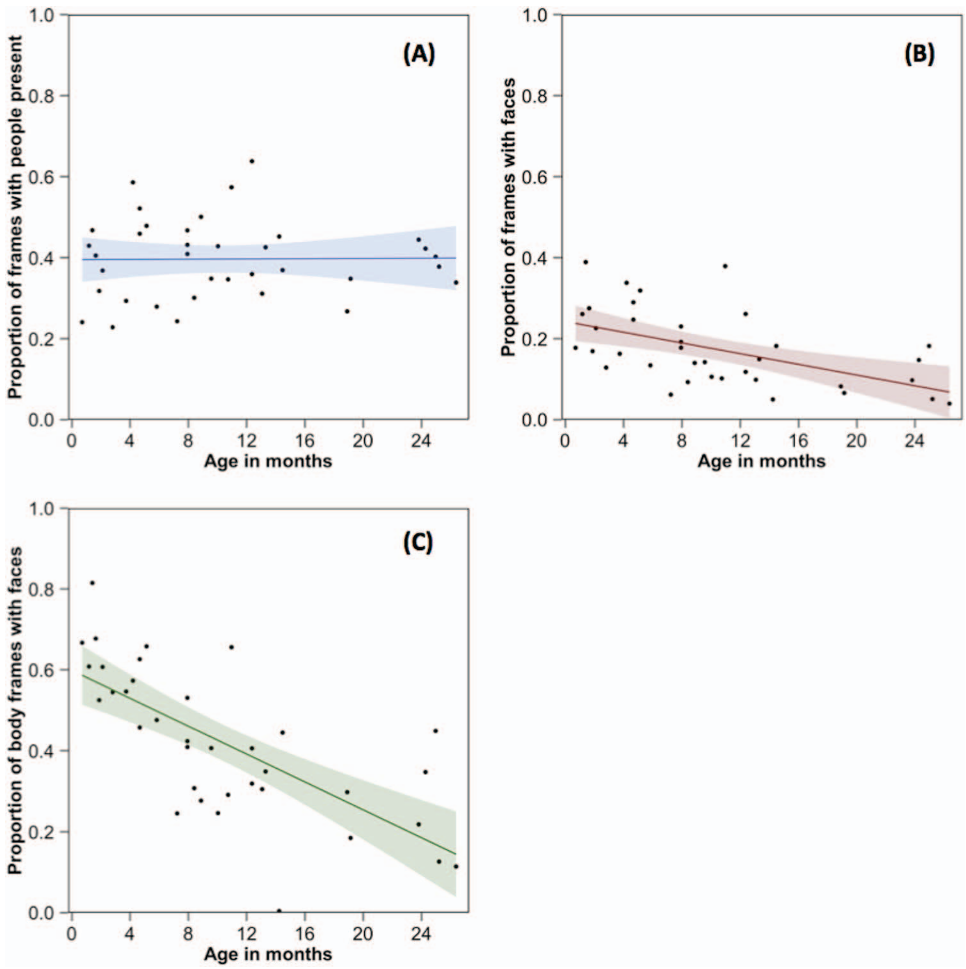


Figure 4. Proportion of total frames with people in view. (A) Proportion of frames that contain at least one person (whole or body part), (B) Proportion of frames that contain at least once face (or face part), (C) Proportion of person frames that also display faces. The bands show 95% confidence intervals around the best-fit line. See the online article for the color version of this figure.

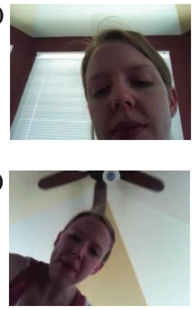
ference in people views of younger and older infants, showing the proportion of people views that include faces declines systematically with age, $R^2 = 0.50$, $F(1, 34) = 34.21$, $p < .001$. For very young infants, people information is highly selective and more often includes a face (65% of person views contain faces for infants 2 months or younger), but faces are much more rarely part of the people information for older infants (25% for infants over 18 months). In sum, the evidence is clearly consistent with the face-input hypothesis. It is faces not people that are especially frequent in younger infants' visual experiences and face frequency that declines with age.

Are the faces in the head camera images of the youngest infants also close? This is a critical question because the visual system of infants younger than 3 months, particularly visual acuity, is not as developed as that of older infants (Dobson, Teller, & Belgum, 1978; Maurer & Lewis, 2001a, 2001b; Tondel & Candy, 2008). Thus, the most visually relevant faces for the youngest infants are likely to be those within a few feet of the infant (Maurer & Lewis, 2001a, 2001b) so as to present face-defining features at low spatial frequencies (see Oruc & Barton, 2010). Figure 5 shows views of faces at 1 foot and 2 feet from the infant, and the mean distance and the proportion of faces in the captured head camera images that were closer than 1 foot and closer than 2 feet from the head camera for the seven infants less than 3 months of age. As is apparent, for these young infants, faces are overwhelmingly (more than 60% of the frames overall for infants under 3 months) closer than 2 feet.

In summary, the findings of Experiment 2 provide strong support for the face-input hypothesis. Only the frequency of faces in the infant-perspective images declines with age, not people more generally.

General Discussion

Past research makes clear that the development of human face processing depends on early face experience (Le Grand, Mondloch, Maurer, & Brent, 2001). Because humans are social animals and because human faces provide important social information, it

(A)
(B)


Age (months)	Mean Distance (feet)	Std. error	Proportion < 1 foot	Proportion < 2 feet
0.8	1.35	0.17	.90	.92
1.2	3.23	0.09	.32	.46
1.4	2.62	0.04	.24	.70
1.6	2.46	0.06	.45	.62
1.8	1.92	0.07	.72	.82
2.0	2.46	0.06	.35	.67
2.8	2.15	0.07	.61	.71

Figure 5. Faces close to the infant. This figure shows sample images from head camera video, each showing a face that was captured at (A) 1 foot and (B) 2 feet from the infant wearing the camera. The table to the right lists the mean distances at which infants younger than 3 months encountered faces in front of them during the course of their recording. The table also lists the proportion of total times when faces were within 1 foot and 2 feet from the infant. Almost all of these infants encountered faces within 2 feet for a majority of the time. The individual featured in these images has provided signed consent for her likeness to be published by the authors. See the online article for the color version of this figure.

seems reasonable to assume that faces are visually prevalent throughout life (e.g., Calder & Young, 2005; Cohen & Cashon, 2001; Cohen Kadosh & Johnson, 2007; Haxby et al., 2000; Kanwisher & Yovel, 2006; McKone et al., 2007; Nelson, 2003; Tarr & Gauthier, 2000; Zhao & Bentin, 2008). However, recent research using head cameras suggests potentially significant changes in the faces in infant views (Fausey et al., 2016; Jayaraman et al., 2015; Sugden et al., 2014). The two experiments reported here were designed to answer the question of whether this age-related decline in faces was about a change in the frequency of *people* positioned in front of the infant or more specifically about the frequency of *faces* in front of the infant. The specific empirical contribution, then, is the quantification of the availability of people's bodies versus their faces in infant visual environments as a function of infant age.

The results of the two experiments show that the frequency of people in the vicinity of the infant and in the head camera images did not decrease over the first 2 years of life. However, and as reported before (Jayaraman et al., 2015), the frequency of faces in the head camera images did decline. An estimation of the minutes per hour (from the image sampling rate and proportion of frames with faces) indicates that for infants under 3 months of age, faces were in view for about 14 min per hour but for infants older than 18 months, faces were present less than 5 min per hour. For infants younger than 3 months of age, these in-view faces were also close to the infant (and thus visually large). Because people (any body part) were present at the same rate for older and younger infants (24 min an hour), this means that young infants were receiving a heavy visual dose of people information that included faces (see Jayaraman et al., 2015) whereas older infants were receiving other kinds of nonface information about people (see Fausey et al., 2016). We consider these findings with respect to two issues: First, what might be the contribution of early dense face experiences to the development of face processing? Second, what drives this developmental pattern and what are the implications of these findings for the development of social behaviors?

Face Processing

The development of human face processing shows a long developmental course that is not fully mature until adolescence (see Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis, & Morton, 1991) and the influence of continuing experiences with faces is evident throughout infancy (Scott & Monesson, 2010). In particular, previous research shows that the discrimination of different classes of people changes systematically over the first year of life beginning with a more general competency that can discriminate all kinds of faces to a narrower more specific one that is expert in discriminating the classes of faces (gender, race, age) that have been most common in the infant's environment (Scott & Monesson, 2010). With age and experience, infants and then children also become better able to categorize and interpret facial gestures (Scott & Monesson, 2010; Scott et al., 2007) and their representation of face information becomes more holistic and configural (Carey & Diamond, 1994; Maurer, Mondloch, & Lewis, 2007; Mondloch, Pathman, Maurer, Le Grand, & de Schonen, 2007; Tanaka & Farah, 1993). One could imagine the case in which the frequency of faces is constant across development and this increasing face-expertise with age is the product of aggregated experiences over the day-in and day-out experience of faces. But the evidence suggests that the face-learning environment is not developmentally static and differs markedly, for example, between the 2-month-old and the 24-month-old. Part of the explanation of how and why human face perception develops the way it does seems likely to lie in the developmentally changing frequency and character of the faces in the input.

What is the role of early visual experiences that are particularly dense with faces? According to Gottlieb's theory (Gottlieb, 1991, 2007) of probabilistic epigenesis and experiential canalization, species-important outcomes arise not solely from genetic determination but from the interaction of intrinsic properties of the developing organism and its species-typical environment. From this perspective, dense faces in very early (but not later) visual experience of people might itself be a product of evolutionary selection (Gottlieb, 1991). Research with children who were deprived of early (first 2 to 6 months of life) visual input by congenital cataracts that were subsequently removed supports the idea this dense early visual experience might be critical for later developments in face processing. In particular, early deprivation results in permanent deficits in configural face processing, deficits that only become evident much later in development (Le Grand et al., 2001). Maurer, Mondloch, and Lewis (2007) offered the term "sleeper effects" in their interpretation of the findings and specifically hypothesized that early experiences preserved and established the neural substrate for abilities that develop much later (see also Byrge, Sporns, & Smith, 2014). In brief, dense *early* experiences with faces may be necessary to set up or maintain the cortical circuitry that supports the later emergence of specialized face processing. These early dense experiences may also provide the foundation for later learning from visually less frequent faces. The present results also suggest an alternative to interpretation; that is, it may be the density and kinds of face experience, not its timing early in development, that is critical to the neural circuitry. If caretakers do not interact with postcataract older infants the way they do with very young infants, infants who have their first visual experiences after 3 months, may never have quite the same input, that is, may never experience several months of dense close faces, the density of face experience that may be critical to the development of specialized face processing. This is an important possibility that merits further research.

What Drives the Change in Face but Not People Frequency?

There are two relevant questions to understanding the factors that lead to the observed developmental trend of a decline in the frequency of faces in view, but no decline in the frequency of people in view: What factors encourage or support early face experiences? What factors contribute to the decline of specifically faces but not people? Given the importance of face processing to human social interactions, one would expect strong evolutionary pressures to ensure that faces are available for young infants. One way to do this is to bias very young infants to want to look at faces; another way to do this is to bias mature humans to put their faces in front of very young infants. There is clear evidence for the first bias in very young infants' preference to look at simple "face–like" arrays that consist of two dark blobs (eyes) within a face-shaped contour (Goren et al., 1975; Johnson et al., 1991). Theorists (McKone et al., 2007; Morton & Johnson, 1991) have interpreted this early preference as an indication of the species importance of early exposure to faces since this bias on the part of infants would have the effect of increasing the time faces are in view. However, a bias to look at faces may not be enough by itself and thus evolutionary pressures may have also ensured that people put their faces (and not just their bodies) in front of infants. Very young infants, with their limited motor abilities, have little control over the contents of the scenes in front of their faces: They cannot readily turn their heads or change posture to bring off-side or obstructed faces into clear view. Thus, the selective pressure that makes faces dense in early visual environments may not reside in the infant alone. In his theory of the evolutionary and developmental constraints on human mother–infant attachment, Bowlby (1988) pointed to multiple factors that not only brought caretakers close to their young infants but that specifically fostered face-to-face interactions. These include the social-emotional responses of caretakers to infant gaze directed at the caretaker's face and more generally, caretaker response to infant behaviors contingent on the caretaker's facial gestures and voice. The bidirectional pulls of each social partner on the other during this early period of infancy create enduring periods of faces directly in front of faces. Thus, the main factor responsible for the high frequency of faces in the early visual experiences of infants may lie, as Bowlby proposed, not in the infant alone and not in the caretaker alone, but in the joint constraints that bring them face-to-face and that make those interactions rewarding to both participants. Thus, a critical next step in this program of research is to determine whether the dense early experience of faces is primarily about face-to-face play with it multimodal and dynamic components.

What drives the decrease in the frequency of faces in the infant view after those early months? An answer that is easy to comprehend would have been that parents are less often near the older infant, or less often positioned in front of the older infant. The unique contribution of the present study is the evidence that this easy answer is wrong. People *are* near and visually in view of infants across the full range of infancy. The question that we now know needs to be answered is what drives a decrease in faces in view but not a decrease in the people in view? At present, we do not have an answer. The contribution of head cameras is that they capture the perceivers' egocentric views of the world. These are views with fundamentally different properties than our usual third person views (Fathi et al., 2011; Lee, Ghosh, & Grauman, 2012; Li, Fathi, & Rehg, 2013; Smith et al., 2014) and that they are created by the perceiver's own momentary actions within the perceiver's local context. A growing literature in computational vision and in psychology on egocentric views (Foulsham, Walker, & Kingstone, 2011; Li et al., 2013; Yamada et al., 2011; Yu & Smith, 2013) is making clear that the properties of these views (and thus their potential momentary causes) are often unexpected, challenging in multiple ways what we think we know about visual environments from other measures.

One relevant question is who is responsible for the change. The pattern could lie principally in changes in infants' interests or in parent behavior. For example, parents *could* place their faces just as often in front of older infants but older infants could actively be more interested other aspects of the visual scene and turn their heads and bodies in other directions. Alternatively, older infants could be interested in faces (Frank, Vul, & Saxe, 2012) but parents could interact with older infants differently than they do with younger infants. We suspect that changes on both the parent side and the infant side of the equation matter. Relevant to this idea, there is a large literature documenting a shift from the face-to-face social interactions in early infancy to the more object centered (or triadic) interactions in later infancy (Bakeman & Adamson, 1984; Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998; Cleveland & Striano, 2007). This shift in social interactions may bring different body parts, and particularly hands acting on objects, into view (see Fausey et al., 2016; Yu & Smith, 2013). Likewise, language and the growth of social interactions at a distance through words (see Hart & Risley, 1999) seems likely to change children's views of people.

One limitation of the present study is that it included families from just one cultural context. Parenting practices are known to vary widely across different cultures (Bornstein, 1995). We do not know the impact of these different practices on infant visual experience nor the potential impact of cultural differences in visual experiences on the development of visual face processing. We do know that congenital cataracts, even when corrected when infants were 3 to 6 months of age (Maurer et al., 2007) and institutionalization (with reduced early face experiences, Moulson, Westerlund, Fox, Zeanah, & Nelson, 2009) have lasting effects on face processing. These findings are consistent with the proposal that early dense visual experiences of faces is central to the development of human face processing and thus possibly universal across cultures. However, there are also well-documented cultural effects in nonface visual processing, in adults (Caparos et al., 2012; Chua, Boland, & Nisbett, 2005; Han & Northoff, 2008) and in children (Kuwabara & Smith, 2012, in press). A key theoretical question is whether face experiences and the development of face processing are immune to these cultural effects on visual experience and visual development.

In conclusion, a complete understanding of the role of experience in the development of specialized face processing, and visual development more generally, requires an understanding of infants' visual experiences and how those experiences change with development. At present, we know very little. The current study provides new evidence relevant to the factors that underlie the dense presence of faces in early visual experience and the later decline in the frequency of faces in infants' views. From the present study, we know that the relevant factors, contexts, and behaviors that give rise to the age-related decline in the frequency of in-view faces are specific to faces. Therefore, the age-related decline in the frequency of faces in infant visual experience cannot be explained by the changes in the number people in the same vicinity as the infant nor even by the number of people in the infant's egocentric view.

References

Aslin, R. N. (1981). Experiential influences and sensitive periods in perceptual development: A unified model. *Development of Perception, 1,* 5–44.

Aslin, R. N. (2009). How infants view natural scenes gathered from a head-mounted camera. *Optometry and Vision Science, 86,* 561.

Bakeman, R. (1997). *Observing interaction: An introduction to sequential analysis*. Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511527685

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development, 55,* 1278–1289. http://dx.doi.org/10.2307/1129997

Ballard, D. H., Hayhoe, M. M., Li, F., Whitehead, S. D., Frisby, J., Taylor, J., & Fisher, R. (1992). Hand-eye coordination during sequential tasks. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 337,* 331–338. http://dx.doi.org/10.1098/rstb.1992.0111

Bambach, S., Crandall, D. J., & Yu, C. (2013). *Understanding embodied visual attention in child-parent interaction*. Paper presented at the Development and Learning and Epigenetic Robotics (ICDL), 2013 IEEE Third Joint International Conference on.

Bertenthal, B. I., & Campos, J. J. (1990). A systems approach to the organizing effects of self-produced locomotion during infancy. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (pp. 1–60). Norwood, NJ: Ablex.

Bloch, H., & Carchon, I. (1992). On the onset of eye-head coordination in infants. *Behavioural Brain Research, 49,* 85–90. http://dx.doi.org/10.1016/S0166-4328(05)80197-4

Bornstein, M. H. (1995). Form and function: Implications for studies of culture and human development. *Culture and Psychology, 1,* 123–137. http://dx.doi.org/10.1177/1354067X9511009

Bowlby, J. (1988). Attachment, communication, and the therapeutic process. A secure base: Parent-child attachment and healthy human development, 137–157.

Braddick, O., & Atkinson, J. (2011). Development of human visual function. *Vision research, 51,* 1588–1609.

Bushnell, W. R. (2003). Newborn face recognition. In O. Pascalis & A. Slater (Eds.), *The development of face processing in infancy and early childhood: Current perspectives* (pp. 41–53). Hauppauge, NY: Nova Science Publishers, Inc.

Byrge, L., Sporns, O., & Smith, L. B. (2014). Developmental process emerges from extended brain-body-behavior networks. *Trends in Cognitive Sciences, 18,* 395–403. http://dx.doi.org/10.1016/j.tics.2014.04.010

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience, 6,* 641–651. http://dx.doi.org/10.1038/nrn1724

Caparos, S., Ahmed, L., Bremner, A. J., de Fockert, J. W., Linnell, K. J., & Davidoff, J. (2012). Exposure to an urban environment alters the local bias of a remote culture. *Cognition, 122,* 80–85. http://dx.doi.org/10.1016/j.cognition.2011.08.013

Carey, S., & Diamond, R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition, 1,* 253–274. http://dx.doi.org/10.1080/13506289408402302

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63,* 1–143. http://dx.doi.org/10.2307/1166214

Cassia, V. M., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science, 15,* 379–383. http://dx.doi.org/10.1111/j.0956-7976.2004.00688.x

Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 12629–12633. http://dx.doi.org/10.1073/pnas.0506162102

Cleveland, A., & Striano, T. (2007). The effects of joint attention on object processing in 4- and 9-month-old infants. *Infant Behavior and Development, 30,* 499–504. http://dx.doi.org/10.1016/j.infbeh.2006.10.009

Cohen, L. B., & Cashon, C. H. (2001). Do 7-month-old infants process independent features or facial configurations? *Infant and Child Development, 10,* 83–92. http://dx.doi.org/10.1002/icd.250

Cohen Kadosh, K., & Johnson, M. H. (2007). Developing a cortex specialized for face perception. *Trends in Cognitive Sciences, 11,* 367–369. http://dx.doi.org/10.1016/j.tics.2007.06.007

Daniel, B. M., & Lee, D. N. (1990). Development of looking with head and eyes. *Journal of Experimental Child Psychology, 50,* 200–216. http://dx.doi.org/10.1016/0022-0965(90)90039-B

Dobson, V., Teller, D. Y., & Belgum, J. (1978). Visual acuity in human infants assessed with stationary stripes and phase-alternated checkerboards. *Vision Research, 18,* 1233–1238.

Fantz, R. L. (1963). Pattern vision in newborn infants. *Science, 140,* 296–297. http://dx.doi.org/10.1126/science.140.3564.296

Fathi, A., Ren, X., & Rehg, J. M. (2011, June). Learning to recognize objects in egocentric activities. In *IEEE Conference On Computer Vision and Pattern Recognition (CVPR), 2011* (pp. 3281–3288). Colorado Springs, CO: IEEE.

Fausey, C. M., Jayaraman, S., & Smith, L. B. (2016). From faces to hands: Changing visual input in the first two years. *Cognition, 152,* 101–107. http://dx.doi.org/10.1016/j.cognition.2016.03.005

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision*

Research, 51, 1920–1931. http://dx.doi.org/10.1016/j.visres.2011.07.002

Franchak, J. M., & Adolph, K. E. (2012). What infants know and what they do: Perceiving possibilities for walking through openings. *Developmental Psychology, 48,* 1254–1261. http://dx.doi.org/10.1037/a0027530

Frank, M. C., Simmons, K., Yurovsky, D., & Pusiol, G. (2013). Developmental and postural changes in children's visual access to faces. In *Proceedings of the 35th annual meeting of the Cognitive Science Society* (pp. 454–459). Austin, TX: Cognitive Science Society.

Frank, M. C., Vul, E., & Saxe, R. (2012). Measuring the development of social attention using free-viewing. *Infancy, 17,* 355–375. http://dx.doi.org/10.1111/j.1532-7078.2011.00086.x

Gilmore, R. O., Raudies, F., & Jayaraman, S. (2015, August). What accounts for developmental shifts in optic flow sensitivity? In *2015 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* (pp. 19–25). Providence, RI: IEEE.

Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics, 56,* 544–549.

Gottlieb, G. (1991). Experiential canalization of behavioral development: Theory. *Developmental Psychology, 27,* 4–13. http://dx.doi.org/10.1037/0012-1649.27.1.4

Gottlieb, G. (2007). Probabilistic epigenesis. *Developmental Science, 10,* 1–11. http://dx.doi.org/10.1111/j.1467-7687.2007.00556.x

Han, S., & Northoff, G. (2008). Culture-sensitive neural substrates of human cognition: A transcultural neuroimaging approach. *Nature Reviews Neuroscience, 9,* 646–654. http://dx.doi.org/10.1038/nrn2456

Hart, B., & Risley, T. R. (1999). *The social world of children: Learning to talk*. Washington, DC: ERIC.

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

Hofferth, S. L. (1996). Child care in the United States today. *The Future of Children, 6,* 41–61. http://dx.doi.org/10.2307/1602418

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2015). The faces in infant-perspective scenes change over the first year of life. *PLoS ONE, 10,* e0123780. http://dx.doi.org/10.1371/journal.pone.0123780

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40,* 1–19. http://dx.doi.org/10.1016/0010-0277(91)90045-6

Johnson, M. H., & Morton, J. (1991). *Biology and cognitive development: The case of face recognition*. New York, NY: Blackwell.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 361,* 2109–2128. http://dx.doi.org/10.1098/rstb.2006.1934

Kreader, J. L., Ferguson, D., & Lawrence, S. (2005). *Infant and toddler child care arrangements*. New York, NY: Child Care & Early Education Research Connections.

Kretch, K. S., Franchak, J. M., & Adolph, K. E. (2014). Crawling and walking infants see the world differently. *Child Development, 85,* 1503–1518. http://dx.doi.org/10.1111/cdev.12206

Kuwabara, M., & Smith, L. B. (2012). Cross-cultural differences in cognitive development: Attention to relations and objects. *Journal of Experimental Child Psychology, 113,* 20–35. http://dx.doi.org/10.1016/j.jecp.2012.04.009

Kuwabara, M., & Smith, L. B. (in press). Cultural differences in object recognition in 3-year olds children. *Journal of Experimental Child Psychology*.

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

Lee, Y. J., Ghosh, J., & Grauman, K. (2012, June). Discovering important people and objects for egocentric video summarization. In *IEEE Conference On Computer Vision and Pattern Recognition (CVPR), 2012* (Vol. 2, p. 7). Providence, RI: IEEE.

Li, Y., Fathi, A., & Rehg, J. (2013, December). Learning to predict gaze in egocentric video. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 3216–3223). Sydney, Australia: IEEE.

Maurer, D., & Lewis, T. L. (2001a). Visual acuity and spatial contrast sensitivity: Normal development and underlying mechanisms. *Handbook of developmental cognitive neuroscience,* 237–250.

Maurer, D., & Lewis, T. L. (2001b). Visual acuity: The role of visual input in inducing postnatal change. *Clinical Neuroscience Research, 1,* 239–247.

Maurer, D., Mondloch, C. J., & Lewis, T. L. (2007). Sleeper effects. *Developmental Science, 10,* 40–47. http://dx.doi.org/10.1111/j.1467-7687.2007.00562.x

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences, 11,* 8–15. http://dx.doi.org/10.1016/j.tics.2006.11.002

Mondloch, C. J., Pathman, T., Maurer, D., Le Grand, R., & de Schonen, S. (2007). The composite face effect in six-year-old children: Evidence of adult-like holistic face processing. *Visual Cognition, 15,* 564–577. http://dx.doi.org/10.1080/13506280600859383

Morrissey, T. W., & Banghart, P. (2007). *Family child care in the United States*. New York, NY: Columbia University Academic Commons.

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review, 98,* 164–181. http://dx.doi.org/10.1037/0033-295X.98.2.164

Moulson, M. C., Westerlund, A., Fox, N. A., Zeanah, C. H., & Nelson, C. A. (2009). The effects of early experience on face recognition: An event-related potential study of institutionalized children in Romania. *Child Development, 80,* 1039–1056. http://dx.doi.org/10.1111/j.1467-8624.2009.01315.x

Nelson, C. A. (2003). Can we develop a neurobiological model of human social-emotional development? Integrative thoughts on the effects of separation on parent–child interactions. *Annals of the New York Academy of Sciences, 1008,* 48–54. http://dx.doi.org/10.1196/annals.1301.005

Oruc, I., & Barton, J. J. S. (2010). Critical frequencies in the perception of letters, faces, and novel shapes: Evidence for limited scale invariance for faces. *Journal of Vision, 10,* 20. http://dx.doi.org/10.1167/10.12.20

Pascalis, O., Loevenbruck, H., Quinn, P. C., Kandel, S., Tanaka, J. W., & Lee, K. (2014). On the links among face processing, language processing, and narrowing during development. *Child Development Perspectives, 8,* 65–70. http://dx.doi.org/10.1111/cdep.12064

Pereira, A. F., Smith, L. B., & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin & Review, 21,* 178–185. http://dx.doi.org/10.3758/s13423-013-0466-4

Pirsiavash, H., & Ramanan, D. (2012, June). Detecting activities of daily living in first-person camera views. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on* (pp. 2847–2854). Providence, RI: IEEE.

Ruff, H. A., & Lawson, K. R. (1990). Development of sustained, focused attention in young children during free play. *Developmental Psychology, 26,* 85–93. http://dx.doi.org/10.1037/0012-1649.26.1.85

Scherf, K. S., & Scott, L. S. (2012). Connecting developmental trajectories: Biases in face processing from infancy to adulthood. *Developmental Psychobiology, 54,* 643–663. http://dx.doi.org/10.1002/dev.21013

Schmitow, C., & Stenberg, G. (2015). What aspects of others' behaviors do infants attend to in live situations? *Infant Behavior and Development, 40,* 173–182. http://dx.doi.org/10.1016/j.infbeh.2015.04.002

Schmitow, C., Stenberg, G., Billard, A., & von Hoffsten, C. (2013a). Measuring direction of looking with a head-mounted camera. *International Journal of Behavioral Development, 37,* 471–477. http://dx.doi.org/10.1177/0165025413495749

Schmitow, C., Stenberg, G., Billard, A., & von Hofsten, C. (2013b). Using a head-mounted camera to infer attention direction. *International Journal of Behavioral Development, 37,* 468–474. http://dx.doi.org/10.1177/0165025413495749

Scott, L. S., & Monesson, A. (2010). Experience-dependent neural specialization during infancy. *Neuropsychologia, 48,* 1857–1861. http://dx.doi.org/10.1016/j.neuropsychologia.2010.02.008

Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science, 16,* 197–201. http://dx.doi.org/10.1111/j.1467-8721.2007.00503.x

Smith, L., Yu, C., Yoshida, H., & Fausey, C. M. (2014). Contributions of head-mounted cameras to studying the visual environments of infants and young children. *Journal of Cognition and Development, 16,* 407–419.

Sugden, N. A., Mohamed-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. http://dx.doi.org/10.1002/dev.21183

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology, 46,* 225–245. http://dx.doi.org/10.1080/14640749308401045

Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience, 3,* 764–769. http://dx.doi.org/10.1038/77666

Tondel, G. M., & Candy, T. R. (2008). Accommodation and vergence latencies in human infants. *Vision research, 48,* 564–576.

Wooding, A. R., Boyd, J., & Geddis, D. C. (1990). Sleep patterns of New Zealand infants during the first 12 months of life. *Journal of Paediatrics and Child Health, 26,* 85–88. http://dx.doi.org/10.1111/j.1440-1754.1990.tb02392.x

Yamada, K., Sugano, Y., Okabe, T., Sato, Y., Sugimoto, A., & Hiraki, K. (2011). Attention prediction in egocentric video using motion and visual saliency. In Y.-S. Ho (Ed.), *Advances in image and video technology* (pp. 277–288). New York, NY: Springer.

Yoshida, H., & Smith, L. B. (2008). What's in view for toddlers? Using a head camera to study visual experience. *Infancy, 13,* 229–248. http://dx.doi.org/10.1080/15250000802004437

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition, 125,* 244–262. http://dx.doi.org/10.1016/j.cognition.2012.06.016

Yu, C., & Smith, L. B. (2013). Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PLoS ONE, 8,* e79659. http://dx.doi.org/10.1371/journal.pone.0079659

Zhao, L., & Bentin, S. (2008). Own- and other-race categorization of faces by race, gender, and age. *Psychonomic Bulletin & Review, 15,* 1093–1099. http://dx.doi.org/10.3758/PBR.15.6.1093

Received November 12, 2015
Revision received March 3, 2016
Accepted August 22, 2016 ■