Journal of Experimental Child Psychology 118 (2014) 13–26

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Visual search and attention to faces during early infancy

Michael C. Frank [a,*], Dima Amso [b], Scott P. Johnson [c]

[a] Department of Psychology, Stanford University, Stanford, CA 94305, USA
[b] Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI 02912, USA
[c] Department of Psychology, University of California, Los Angeles, CA 90095, USA

ARTICLE INFO

Article history:
Received 19 July 2013
Revised 28 August 2013

Keywords:
Infancy
Eye tracking
Social attention
Visual search
Face perception
Visual attention

ABSTRACT

Newborn babies look preferentially at faces and face-like displays, yet over the course of their first year much changes about both the way infants process visual stimuli and how they allocate their attention to the social world. Despite this initial preference for faces in restricted contexts, the amount that infants look at faces increases considerably during the first year. Is this development related to changes in attentional orienting abilities? We explored this possibility by showing 3-, 6-, and 9-month-olds engaging animated and live-action videos of social stimuli and also measuring their visual search performance with both moving and static search displays. Replicating previous findings, looking at faces increased with age; in addition, the amount of looking at faces was strongly related to the youngest infants' performance in visual search. These results suggest that infants' attentional abilities may be an important factor in facilitating their social attention early in development.

Introduction

How do infants and young children see the social world? From immediately after their birth, infants attend preferentially to faces and face-like configurations (Farroni et al., 2005; Johnson, Dziurawiec, Ellis, & Morton, 1991). Over the course of their first year, their representations of faces become specific to their particular environment (Kelly et al., 2007; Pascalis et al., 2005), and they begin to be able to make inferences about other agents' internal states such as their goals (Gergely

* Corresponding author.
E-mail address: mcfrank@stanford.edu (M.C. Frank).

http://dx.doi.org/10.1016/j.jecp.2013.08.012

& Csibra, 2003) and their focus of attention (Scaife & Bruner, 1975). Infants recognize other social actors by a wide variety of signals, including the presence of facial features such as eyes, their ability to respond contingently, and even their causal abilities (Johnson, Slaughter, & Carey, 1998; Saxe, Tenenbaum, & Carey, 2005). These results and others suggest a picture of infants as both deeply involved in and increasingly knowledgeable about the social world around them.

Less is known about how these abilities are manifest in the complex task of perceiving and processing the world in real time. Most experimental paradigms addressing infants' social abilities use simple schematic stimuli presented repeatedly in isolation—often in infant-controlled paradigms where individual infants get as much time as they need to process a stimulus. These methods produce reliable results and allow for the measurement of subtle contrasts between conditions, but they do not tell us how effective infants are at using their knowledge in real-time perception (Aslin, 2009; Richards, 2010).

Our previous work used eye-tracking data from infants' viewing of videos to begin to address this question. Frank, Vul, and Johnson (2009) showed 3-, 6-, and 9-month-old infants a set of 4-s clips from an animated stimulus (the *Charlie Brown Christmas* movie) and measured the amount of time they spent looking at the faces of the characters. This study found significant increases in fixation time to the faces of the characters between 3 and 9 months of age. This increase was accompanied by increases in the overall similarity of older infants' fixations to one another and decreases in the amount by which their fixations were predicted by the low-level salience of the movies they saw.

Although this study provided evidence for developmental changes in infants' looking at faces in complex scenes, it gave limited insight into the causes of this developmental change. The middle of the first postnatal year is a time of many changes, and changes in social attention could be driven by a wide variety of factors. For example, changes in social preference could emerge as the result of social learning mechanisms. Children might be learning about the information that can be gleaned from the faces of others (e.g., Scaife & Bruner, 1975; Triesch, Teuscher, Deák, & Carlson, 2006; Walden & Ogan, 1988), and this might drive them to sharpen their preference to look to others. In addition, during this period infants are undergoing substantial motoric development; they are learning to reach for objects and sit unattended, and they are even beginning to crawl. There is growing evidence that these motoric changes may be related to infants' visual preferences (Cashon, Ha, Allen, & Barna, 2013; Libertus & Needham, 2011). Finally, there are many substantial changes in children's visual attention over the period from 3 to 9 months of age (Amso & Johnson, 2008; Colombo, 2001; Dannemiller, 2005; Richards, 2010).

Although it is likely that all of these changes have an impact on children's social attention, in our current work we focused on changes in visual attention. In the Frank et al. (2009) study described above, overall visual salience appeared to pull the youngest infants' attention away from social targets and toward other parts of the stimulus background. We were interested in whether this impression was correct. If developmental change in looking at faces is related to infants' changing attentional abilities, then measures of attentional ability should be expected to correlate with face looking. We employed this logic in our study, although we note that the presence of a correlation between these two measures does not imply a causal relationship. Such a correlation might be driven by independent development because both face looking and visual search are known to undergo developmental changes during the first year of life or might be the product of a third causal factor. We begin to address this issue by controlling for chronological age in our analyses, but we return to the problem of causal inference in the Discussion section.

Visual attention involves a variety of distinct abilities. Following the conceptual framework in Colombo (2001), we can separate baseline alertness, spatial orienting, feature-based attention, and endogenous (sustained) attention to a target. Alertness refers to the simple fact of being awake and able to process stimuli; spatial orienting and feature-based attention deal with finding and recognizing visual stimuli, respectively; and endogenous or sustained attention refers to the ability to maintain focus on a target stimulus. Although understanding how infants identify faces is an important challenge (Johnson et al., 1991; Pascalis et al., 2005; Turati, Valenza, Leo, & Simion, 2005), to answer our questions about social attention, we were primarily interested in how infants orient to and sustain attention to faces in complex scenes.

A group of new studies provide important evidence on this question. Gliga, Elsabbagh, Andravizou, and Johnson (2009), Di Giorgio, Turati, Altoè, and Simion (2012), and Gluckman and Johnson (2013) all showed infants circular displays containing a face and three to five distracter objects. In all three studies, 6-month-olds looked longer at the face compared with the distracters, but the 3-month-olds tested by Di Giorgio and colleagues did not. Results from the first fixation, reflecting early orienting responses, were more mixed; Gliga and colleagues' study and Gluckman and Johnson's study found that 6-month-olds' first fixations were directed toward faces (and toward body parts and animals in the Gluckman and Johnson experiment) more often than chance, but this result was not replicated in Di Giorgio and colleagues' study, perhaps because of the use of less salient grayscale stimuli. Libertus and Needham (2011) used a similar paradigm but with a two-alternative (face vs. toy) presentation and found that whereas 3-month-olds failed to show either fast orienting or sustained attention to the face over the toy, 5-month-olds showed both. Escudero, Robbins, and Johnson (2013) used a similar face versus toy paradigm with 4- and 5-month-olds and found evidence for sustained attentional preferences for faces (but did not measure first fixation), and DeNicola, Holt, Lambert, and Cashon (2013) found evidence for sustained preference but not first fixation in a heterogeneous group of 4- to 8-month-olds. Finally, Gluckman and Johnson (2013) reported sustained preference for faces, body parts, and animals in 6-month-olds, relative to foil stimuli, in addition to more first fixations.

Thus, the evidence on sustained attention to faces is consistent across studies; whereas 3-month-olds do not prefer faces in either dynamic displays or static stimulus arrays, older children show a clear face preference. In contrast, the evidence on face orienting is more mixed, with some studies finding a first fixation face preference in infants older than 5 months and others not finding this preference. Both of these sets of results are compatible with developmental changes in orienting and sustained attention. One possibility is that as younger infants scan the visual world, their attention could be captured by salient visual features of non-face visual stimuli, leading them to attend to these stimuli rather than to faces (a change in orienting ability). This explanation would have the benefit of explaining why first fixation biases are present in some paradigms and with some distracter stimuli but not others. The other possibility is that younger infants' attention to faces slips away more quickly than that of older infants (a change in sustained attention).

In our current study, we examined both of these hypotheses by designing an individual differences paradigm in which 3-, 6-, and 9-month-old infants participated in both a face looking measure and a visual attention measure. To measure developmental changes in visual attention, we chose a simple visual oddball search paradigm in which the infant must find a target that varies in its motion or orientation properties from an array of identical distracters (Amso & Johnson, 2006; Dannemiller, 1998, 2005). Search tasks measure orienting responses rather than sustained stimulus attention (Colombo, 2001), reflecting our primary hypothesis that finding faces in complex displays with salient attention-capturing alternatives may be the problem for young infants (Frank et al., 2009). Nevertheless, we also provide some analysis of sustained attention to faces.

To measure attention to faces, we used dwell time on faces in complex dynamic displays. Building on our previous work, infants in our study watched two different videos: *Charlie Brown,* as in our previous study, and a live-action clip from *Sesame Street.* We selected the 3- to 9-month age range to span the developmental changes in attentional abilities and face representation explored in the previous work in this literature (Amso & Johnson, 2006, 2008; DeNicola et al., 2013; Frank et al., 2009). Consistent with previous work, we predicted developmental increases in looking at faces (Frank et al., 2009) and increases in orienting to targets in the search displays (Dannemiller, 2000, 2005). Although both of these sets of tasks have been used in isolation, to our knowledge no previous study has examined the relation between them. The contribution of our current study is to fill this gap.

Method

Participants

Our target sample was composed of participants at 3, 6, and 9 months of age. To achieve this sample, we recruited 70 infants between 2.5 and 9.5 months of age to participate in our study (3 months,

$n = 35$; 6 months, $n = 16$; 9 months, $n = 19$). Of this group, we excluded those participants who fit any of the following exclusion criteria (many infants fit several of these criteria):

1. They did not complete the visual search ($n = 2$) or free-viewing tasks ($n = 11$).
2. Their calibration could not be adjusted offline to ensure spatial accuracy in the free-viewing task ($n = 6$).
3. They contributed less than 30 s of usable data from the free-viewing videos ($n = 6$).

All exclusion parameters were chosen without reference to study results. The final sample in our study was 23 3-month-olds ($M = 3.0$ months, range = 2.5–3.5, 12 boys and 11 girls), 14 6-month-olds ($M = 5.9$ months, range = 5.4–6.6, 8 boys and 6 girls), and 15 9-month-olds ($M = 8.9$ months, range = 8.5–9.3, 5 boys and 10 girls). The total sample size was 52 infants, for an exclusion rate of 26%. All infants excluded for calibration issues in the final sample were 3-month-olds.

Stimuli

Frames from stimulus materials are shown in Fig. 1. The visual search task was as described in Amso and Johnson (2006). Participants viewed displays with a set of 27 static vertical red rectangles on a black field, with one red target rectangle that varied on either its orientation (24 trials) or motion (24 trials). Orientation trials contained targets that were oriented at 30°, 60°, or 90° from vertical; motion trials contained targets that moved from side to side at speeds of 1, 1.5, or 2 Hz. All trials were presented in random order. Trials began with a moving central fixation point; when infants fixated this point, the search display was presented and remained on screen until either infants fixated the target for a cumulative total of 100 ms (within a 30-pixel radius around the target) or 4 s had elapsed.

The free-viewing task consisted of a 120-s clip of the audio and video from the *Charlie Brown Christmas* movie (an engaging animated film for children) and a 128-s clip, again containing both audio and video, from the children's television program *Sesame Street*. The *Charlie Brown* segment consisted of dialog between several animated children (as well as a short passage with an animated dog); the *Sesame Street* clip contained an opening sequence consisting of children playing with musical background (~50 s), a panning street scene that gradually zoomed in on an adult actor (~35 s), and a static conversation between the actor and two puppets (~40 s). Both stimulus items included many examples of intersensory redundancy (e.g., a mouth moving, synchronized with speech). The free-viewing experiment included two instances of an offline calibration stimulus, which consisted of a brightly colored precessing annulus that moved to nine points arranged in a grid around the display. The order of the *Charlie Brown* and *Sesame Street* videos was randomized across children.

Procedure

Participants visited the laboratory for a single testing session. All infants completed the visual search task (implemented using E-Prime), followed by the free-viewing task (presented using Tobii Clearview software). We chose this consistent task ordering because pilot testing revealed that infants

Fig. 1. Stimuli from the tasks in our experiment. (A) An example display from the static visual search with the target in a 60° trial. (B) An example frame from the introduction portion of the *Sesame Street* stimulus. (C) An example frame from the dialog portion of the *Sesame Street* stimulus. (D) An example frame from the *Charlie Brown* stimulus.

preferred the free-viewing tasks. Had we counterbalanced the order, we would have introduced substantial variance in the visual search task (and also increased our dropout rate), depending on whether visual search was presented earlier or later in the session.

All data collection was done using a Tobii ET-1750 corneal reflection eye tracker operating at 50 Hz. Infants sat in a parent's lap during testing, and parents were asked to look down (away from the screen).

Preliminary data analysis

We performed an offline check and adjustment of eye-tracking calibrations in the free-viewing task using the procedure described in Frank, Vul, and Saxe (2012). We first extracted eye-tracking data for the calibration check stimulus for each infant and used a robust regression algorithm to adjust the data so that the average point of gaze corresponded to the known location of the check stimulus. A human observer then hand-coded whether the algorithm had succeeded. The algorithm was judged to have succeeded if the adjusted eye-tracking data generally appeared within the bounds of the calibration check stimuli and did not show excessive spread, jitter, or drift between the two check stimuli.[1]

For the visual search task, details of the analysis were as in Amso and Johnson (2006). We measured reaction time as the time to fixate the region in which the target appeared, and we excluded trials with reaction times under 200 ms (1.6% of total data) because fixation to the target in these trials was unlikely to be due to stimulus-guided saccades; instead, infants were likely fixating the target by chance. We then calculated accuracy as the portion of non-excluded trials within which infants reached the target before 4 s had elapsed and the trial ended. For purposes of the current analysis, we computed average reaction time and average accuracy for each condition (static vs. moving targets). Participants contributed an average of 6.4 trials (SD = 3.5) of reaction time data in the challenging static condition and an average of 14.6 trials (SD = 6.9) of reaction time data in the moving condition.

In the free-viewing tasks, details of analysis were as in Frank et al. (2009). We first annotated each frame of each of the two videos, noting the bounding box of each face in the frame (including the faces of Snoopy, an animated dog, and the *Sesame Street* muppets Bert and Ernie). We smoothed these regions with a 20-pixel radius so as to avoid excluding gaze that was on the boundary of a face or was outside the face due to inaccuracies in the tracking procedure. We then used calibration-corrected point-of-gaze data to compute the proportion of all recorded gaze that fell within the annotated face regions for each video (using dwell time rather than fixations and saccades).

To analyze infants' sustained attention to faces in the free-viewing data, we extracted all sequences of gaze measurements that fell within a face region consistently for more than 100 ms. For each participant, we then computed the average face fixation sequence length. Note that, in principle, these sequences could consist of multiple fixations within a face, for example, to the eyes and then the mouth. Given the temporal and spatial precision of our data, we did not believe that we could identify individual fixations reliably within such small targets.

Not all infants contributed data for both free-viewing videos, although the majority did (83%). Infants contributed gaze data for a mean of 60% of the *Charlie Brown* video and 70% of the *Sesame Street* video. Missing data were due to child motion, blinks, loss of interest, and loss of eye track due to technical issues. Overall, participants contributed a mean of 161 s of total eye-tracking data in the free-viewing tasks (range = 36–244 s). All correlational analyses reported below were conducted over participant averages, with each participant's performance being averaged across the data contributed by that participant (without interpolation).

[1] Because of the inherent subjectivity of this checking operation, we created two thresholds: one strict and one permissive. The permissive threshold required only a relatively general correspondence between eye-tracking data and the calibration points and resulted in six excluded calibrations overall. We report data from this threshold value in the Participants and Results sections. The strict threshold required a much more limited spread of the data and no evidence for any drift in calibration across the experiment. This criterion led to the exclusion of an additional 11 infants, for a total sample of 15 3-month-olds, 13 6-month-olds, and 13 9-month-olds. Where appropriate, we report in footnotes the results of our analyses when carried out using the strict criterion.

Results

We first present the results from the visual search and free-viewing tasks independently. We then present interrelations between these tasks.

Individual task analyses

Both of the tasks we examined showed the same patterns as had been observed in previous reports (Amso & Johnson, 2006; Frank et al., 2009). Each measure is plotted against the age of the participants in Fig. 2.

In the visual search task, looking to the target was quicker and more accurate in the moving target condition compared with the static target condition (moving accuracy = 76.3%, reaction time [RT] = 1371 ms; static accuracy = 33.6%, RT = 1686 ms). In both conditions, accuracy was positively correlated with age (moving: $r = .70$, 95% confidence interval [*CI*] = .52 to .81, $p < .0001$; static: $r = .42$, 95% *CI* = .17 to .62, $p = .002$). Reaction times were also negatively correlated with age in the

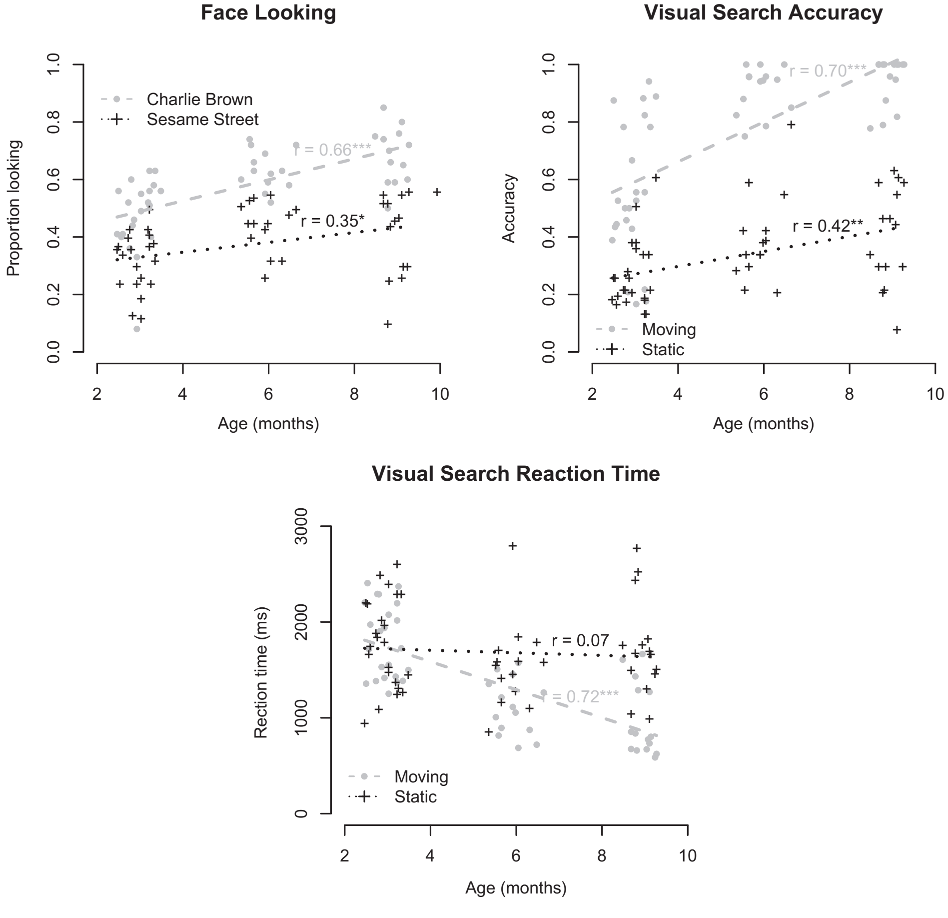

Fig. 2. Proportion looking at faces (left top), visual search accuracy (right top), and visual search reaction time (bottom), plotted by age in months. Dashed/dotted lines and crosses/open circles show individual participants and line of best fit for each condition (*Charlie Brown* vs. *Sesame Street* or static vs. moving targets), respectively. *$p < .05$; **$p < .01$; ***$p < .001$.

moving condition ($r = -.71$, 95% $CI = -.83$ to .55, $p < .0001$), but there were likely too few correct searches in the static condition to produce accurate measurements; hence, there was no reliable correlation with reaction time ($r = -.07$, 95% $CI = -.34$ to .21, $p = .61$). Overall, these results are congruent with those reported by Amso and Johnson (2006); static search was slower and more difficult than moving search.

Looking to the faces of characters in the *Charlie Brown* and *Sesame Street* videos increased significantly with age of participants ($r = .66$, 95% $CI = .46$ to .80, $p < .0001$, and $r = .35$, 95% $CI = .07$ to .58, $p = .02$, respectively). Overall rates of looking at faces in the *Charlie Brown* stimulus were higher than those in the *Sesame Street* stimulus ($M = 58\%$, $SD = 14$ vs. $M = 37\%$, $SD = 13$). These two tasks were highly correlated with one another ($r = .53$, 95% $CI = .27$ to .72, $p = .0002$) even when controlling for age (Pearson partial correlation $r = .44$, $p = .002$).

Nevertheless, we noticed that there were substantial differences in face looking between the first two parts of the *Sesame Street* stimulus, which featured small faces and considerable motion of both the camera and the people in the film, and the second part, which consisted of a series of larger static faces talking to one another. These differences are plotted in Fig. 3. Looking at faces was close to ceiling in the uncomplicated dialog section and showed no reliable developmental trend, whereas looking was much lower in the more complex introduction and showed a reliable developmental increase.

Looking at faces was a separable measure from overall attention toward the screen. In the *Charlie Brown* stimulus, these measures were not reliably correlated ($r = .17$, 95% $CI = -.06$ to .49, $p = .24$). In the *Sesame Street* stimulus as a whole, there was a substantial correlation ($r = .61$, 95% $CI = .39$ to .76, $p < .0001$), but this correlation was not reliable in the first and second sections of the video independently ($r = .23$, 95% $CI = -.12$ to .44, $p = .11$, and $r = .07$, 95% $CI = -.24$ to .36, $p = .67$). Thus, it seems likely that the overall correlation in the *Sesame Street* stimulus was caused by some infants contributing data only for the first part (leading to low overall looking times and low amounts of face looking in the complex first sections) and some infants watching the entire video (leading to both longer looking times and more face looking in the simpler dialog section). This pattern did not vary reliably across ages groups in a one-way analysis of variance (ANOVA) performed on changes in looking time between parts ($F(2,49) = 1.17$, $p = .32$), suggesting that it is unlikely to be a confound in further age-related analyses.

An additional analysis of the free-viewing data examined the length of bouts of sustained attention to faces. As described above, we found the mean length of face-focused gaze segments for each participant for the two videos. These measures were quite correlated with total amount of looking at faces for both *Charlie Brown* ($r = .41$, 95% $CI = -.14$ to .63, $p = .004$) and *Sesame Street* ($r = .70$, 95% $CI = -.51$ to .82, $p < .001$) stimuli. However, they were not significantly correlated with age ($r = .06$, 95% $CI = -.23$ to .35, $p = .67$, and $r = -.05$, 95% $CI = -.32$ to .24, $p = .76$, respectively).

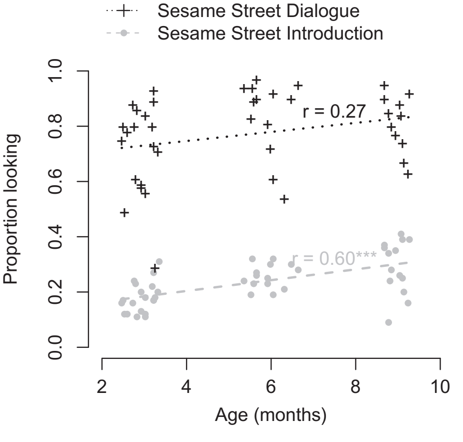

Fig. 3. Proportion looking at faces in the first and second parts of the *Sesame Street* stimulus, plotted by age in months. Dashed/dotted lines and crosses/open circles show individual participants and line of best fit for each condition, respectively. ***$p < .001$.

Because of the strong relation between face looking in the two different stimulus videos, we averaged these two measures to create a composite measure to use in future analyses. This step also ensured that we included all data from all measures available for each infant, thereby maximizing reliability. We also note that there were no significant differences in face looking by gender within any age group (all $ps > .15$ in two-sample t tests).

Taken together, the relation between the *Charlie Brown* and *Sesame Street* face looking measures, their relation to age, and their consistency with prior results all suggest that looking at faces in complex dynamic displays constitutes a relatively stable behavior that develops over the first year of life. These results replicate and extend the findings of Frank et al. (2009), suggesting that it was neither the short clips nor the animated stimuli used in that experiment that led to the observed developmental results.[2]

Relations between visual search and face looking

We next examined the relations between measures of visual search performance and face looking. A full matrix of correlations between the various measures is given in Table 1. We highlight some aspects of this pattern below for further analysis.

Relations between the composite measure of face looking and the visual search variables are shown in Fig. 4. The strongest relation between this composite measure and the search tasks was with accuracy in the moving search condition ($r = .55$, 95% $CI = .33$ to .72, $p < .001$). Accuracy in the static search condition had a weaker relationship ($r = .26$, 95% $CI = -.02$ to .50, $p = .06$). The relation with moving accuracy remained significant when controlling for age (Pearson partial correlation $r = .32$, $p = .02$), but the static search accuracy correlation was no longer reliable (partial $r = .05$, $p = .70$).

Face looking was also significantly correlated with reaction time in both the moving ($r = -.48$, 95% $CI = -.66$ to $-.24$, $p < .001$) and static ($r = -.33$, 95% $CI = -.55$ to $-.06$, $p = .02$) search tasks. The correlation with static reaction time remained significant after controlling for age ($r = -.34$, $p = .01$), whereas there was at most a limited remaining effect on moving reaction time ($r = -.19$, $p = .19$).

The analysis of the relation among face looking, moving search accuracy, and age can also be repeated as a simple mediation analysis within a linear regression framework (Baron & Kenny, 1986). Moving search accuracy alone predicts face looking (= .27, $p < .001$), as does age (in months) alone (= .025, $p < .001$). However, when both predictors are entered into the regression, moving search remains significant (= .19, $p = .02$), whereas age is no longer significant (= .012, $p = .13$). Hence, we are justified in concluding that moving search accuracy mediates the effects of age on face looking. This relation is pictured in Fig. 5. However, we did not find a similar mediation relation for moving reaction time, static reaction time, or accuracy.

Much the same relation between face looking and moving visual search held within the 3-month-old group alone. Face looking was nonsignificantly correlated with age ($r = .35$, 95% $CI = -.07$ to .67, $p = .10$), but moving search accuracy and face looking were reliably correlated ($r = .46$, 95% $CI = .05$ to .73, $p = .03$), and a linear regression including both as predictors left a marginally significant effect of accuracy (= .17, $p = .06$) with no effect of age (= .003, $p = .21$).[3] Within the 6- and 9-month-old age groups, none of these relationships came close to achieving significance.

We repeated a similar set of analyses using the mean length of bouts of attention to faces rather than the total proportion of looking at faces. Only one of these analyses was close to statistical significance; length of looking at *Sesame Street* correlated with static visual search reaction times ($r = -.31$, 95% $CI = -.55$ to $-.03$, $p = .03$). Perhaps infants who were faster at finding targets in the more difficult static search paradigm also sustained attention to faces longer in the more complex *Sesame Street* video. Consistent with this interpretation, the correlation was stronger and close to statistical significance in the 9-month-old group alone ($r = -.46$, 95% $CI = -.79$ to .09, $p = .10$) but not in any of the

2 Using the strict calibration inclusion criterion defined above, our results remain largely unaltered. The pattern of reliabilities for correlations remains, although the correlation between the full *Sesame Street* stimulus and participants' age is reliable at $p = .08$.

3 Using the strict sample defined above, the reliability of correlations between face looking and search tasks remains unchanged. The mediation results for moving search accuracy are close to statistical significance with both the full sample and the 3-month-old group, with no reliable age coefficient.

Table 1
Correlations between measures in the free-viewing and visual search tasks.

	SS Face	CB Attn	SS Attn	Mv Acc	Mv RT	St Acc	St RT
CB Face	.53***	.23	.08	.59***	–.60***	.31*	–.21
SS Face		.13	.61***	.45***	–.40**	.25	–.33*
CB Attn			.22	.06	.02	–.07	–.03
SS Attn				.08	.00	.01	–.10
Mv Acc					–.73***	.48***	–.08
Mv RT						–.53***	.13
St Acc							–.09

Note: CB, *Charlie Brown*; SS, *Sesame Street*; Attn, total looking at stimulus movie; Mv, moving search; St, static search.
* $p < .10$.
** $p < .05$.
*** $p < .001$.

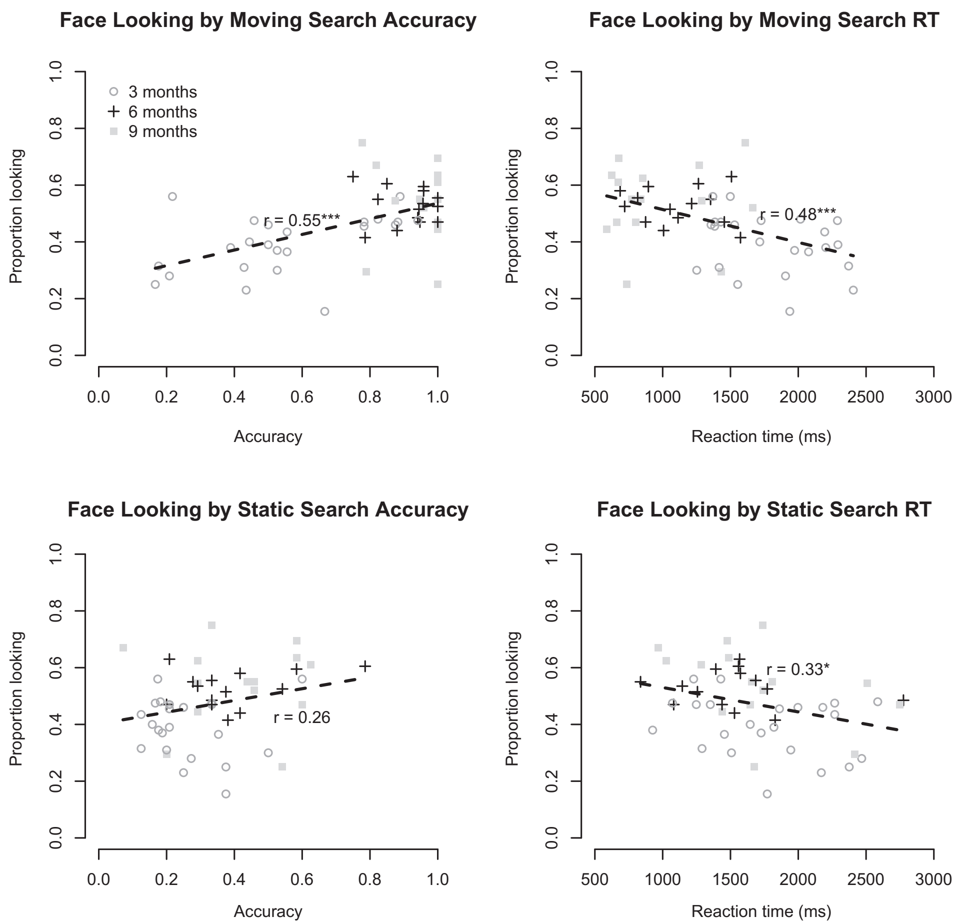

Fig. 4. Average proportion of face looking during *Charlie Brown* and *Sesame Street*, plotted by accuracy (left panels) and reaction time (right panels) in moving (top panels) and static (bottom panels) search tasks for all three age groups. Individual 3-, 6-, and 9-month-olds are shown by open circles, crosses, and filled squares, respectively, with the dashed line showing a line of best fit. $^{*}p < .05$; $^{***}p < .001$.

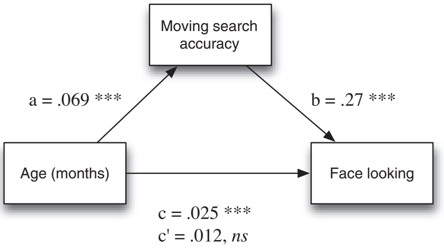

Fig. 5. Mediation relationship among age, moving search accuracy, and a composite measure of face looking. Letters show coefficient weights; *c* indicates the direct relationship between age and face looking; and *c*′ indicates the relationship after controlling for moving search accuracy.

other age groups. However, we interpret this result with caution; it is relatively small in magnitude and would not survive correction for multiple comparisons across the combination of bout length measures and search measures (a total of eight comparisons). Overall, total proportion looking at faces produced more robust and consistent results than length of sustained bouts of attention to faces.

To summarize, we found a strong link between the amount of time that infants looked at faces in the free-viewing task and their accuracy and reaction time in the visual search task. Not only were visual search performance and face looking related to one another, but also this relation was not mediated by infants' age; in fact, if anything, some measures of visual search performance mediated age. This relationship held most strongly within the youngest age group in the study, either because these infants were undergoing the greatest developmental changes or because our search measures were most diagnostic for this group. Overall, these results suggest that looking at faces in complex displays relies on the ability to find them.

Discussion

Even though neonates may often look at face-like images over other visual stimuli in forced-choice and tracking experiments, infants' looking at faces in complex displays increases considerably over the first year. One reason for this change may be the greater attentional wherewithal of older infants. Our data provide support for this hypothesis; infants who showed weaker attentional abilities also looked less at faces. This relation was seen primarily in the youngest infants (the 3-month-old group) and was stronger than the relation between chronological age and face looking (both in that group and in the entire sample). In addition, it seemed to hold primarily for the total amount of looking at faces as opposed to the length of bouts of attention to faces, suggesting that search performance was likely related to finding faces rather than sustaining attention to them. Thus, our data support the claim that attentional abilities play an important part in the social preferences that are manifest during early infancy—especially the first 2 to 4 months. Although very young infants may be motivated in general to look at faces, visual attention is necessary to find them in a complex environment full of other salient and distracting stimuli.

Study limitations

In our study, we used visual search as our measure of attentional abilities. Prior work led us to believe that search abilities might be important for very young infants' social attention; infants in the 2- to 4-month age range show substantial individual and developmental variation in search ability (Amso & Johnson, 2006; Dannemiller, 2005), and salient distracter stimuli differentially attract attention at this age (Frank et al., 2009). Nevertheless, the lack of other attentional measures constitutes a significant limitation of our design. There are likely developmental intercorrelations between many aspects of visual attention, and visual search may simply be one task in which relatively precise within-participants measures of infants' attention are available. Other studies could discover a more fine-grained dependency between aspects of social attention and particular attentional abilities such as sustained attention (Colombo, 2001), attentional tracking (Richards & Holley, 1999), disengagement from salient stimuli, and maintenance of representations across occlusion (Johnson, Amso, & Slemmer, 2003). Thus, further parcellating which attentional abilities best predict social looking behavior is a task for future work.

We saw overall stronger relations between face looking and performance in the moving search compared with the static search. Our data cannot distinguish between two possible interpretations of this result. First, the greater correlations for moving search may simply be due to the greater diagnosticity of the moving search task; static search was difficult for all infants in the sample and may have resulted in floor effects. Second, greater correlations may be due to a more fundamental congruence between moving search and the abilities necessary for finding faces. Faces are often moving and talking, especially in complex scenes, and such intermodal redundancies can be powerful cues for young infants (Bahrick & Lickliter, 2000; Bahrick, Lickliter, & Flom, 2004). Because we did not manipulate the presence of synchronized audio, the current design does not allow us to make inferences about the magnitude of intersensory redundancy effects in driving face looking. Nevertheless, we note that intersensory redundancy is not likely to be the sole source of developmental changes in face looking because such changes are observed even in studies that do not feature synchronized audio (e.g., Di Giorgio et al., 2012; Libertus & Needham, 2011).

Our study used a classic individual differences paradigm in which all infants were tested in a single session, with tasks presented in the same order (e.g., Baron, 1979; Johnson, Davidow, Hall-Haro, & Frank, 2008; Treiman, 1984). Both of these steps could have increased the correlations between tasks; we address each in turn. First, testing individual difference tasks on the same day is a common practice, especially given the practical concerns of running experiments with very young infants. Nevertheless, such a paradigm runs the risk of finding correlations due to the infants' general state on each day. These effects would likely be mediated by arousal factors. To investigate this possibility, we used combined time on task in the free-viewing stimuli as a proxy; infants having "better days" should look at the screen more. Even after controlling for time on task and age in regression analyses, we still find that moving search accuracy and reaction time are both significant predictors of looking at faces ($ps < .02$). Thus, we do not believe that "good day" effects are solely responsible for the correlations we observed.

Second, our experimental design was subject to carryover effects in which some aspect of performance on one task could influence a second task. However, the question of how to design an individual differences study is not straightforward. Using a constant task ordering can risk carryover effects, but the alternative—a counterbalanced task order—also introduces substantial issues; a less engaging task can cause decrements in attention to later tasks if it is tested first in some children. This asymmetry between counterbalance conditions can lead to an increase in variability and, hence, a decrease in statistical power. Indeed, this concern motivated our current design because the visual search task was less engaging than the free-viewing tasks. Nevertheless, an order-counterbalanced design with a larger sample would provide a strong further test of our current results.

Finally, our data do not allow us to address the question of whether young infants prefer faces to other stimuli. A large body of work has suggested that such a preference does exist, although it may be driven by a combination of general stimulus-level biases (e.g., Farroni et al., 2005; Johnson et al., 1991; Macchi, Turati, & Simion, 2004; Simion, Cassia, Turati, & Valenza, 2001). This initial preference then becomes more tightly linked to the specific characteristics of human faces over the course of the first year of life (Pascalis, de Haan, & Nelson, 2002; Pascalis et al., 2005; Turati et al., 2005). But when experiments use complex displays or perceptually salient distracters, a face preference is no longer observed in 3-month-olds even though it is observed again as soon as 1 month later (DeNicola et al., 2013; Di Giorgio et al., 2012; Escudero et al., 2013; Frank et al., 2009; Libertus & Needham, 2011). This temporary dip in preference likely does not pose any significant developmental problem; in natural contexts, faces are still likely to be an overwhelmingly frequent target of infants' fixation. Nevertheless, the dip signals that some aspect of young infants' response to faces is relatively fragile, whether it is their preference per se or their ability to realize this preference in the face of attentional demands.

Future directions

Looking forward, the fragility of face preferences is a clue that we can use to understand the mechanisms driving developmental changes in face looking. We considered three possible sources of the change from previous literature: attentional, social, and motoric changes. Although our data show a correlation between attentional abilities and looking at faces, they do not establish a causal relationship. It will be for future work, perhaps using longitudinal methods, to provide further tests. Our data also do not rule out the influence of other factors; in fact, we believe that these factors are likely to play a role in the development of early social attention. For example, Libertus and Needham (2011) found no face preference in untrained 3-month-olds but found a consistent preference in 3-month-olds who had received manual reaching training. These intriguing results suggest that motor development might be one cause of developmental changes in face preference, although they are probably not the driver of our own results (because presumably our youngest infants were not reaching yet given the standard developmental timing).

Our study raises the question of the importance of visual search skill in infants' natural environment. A number of promising techniques now exist to measure social input in more naturalistic settings, providing the possibility of a precise answer to this question. Head-mounted cameras and eye trackers are beginning to provide detailed measurements of the visual world of infants and young children (Aslin, 2009; Cicchino, Aslin, & Rakison, 2011; Franchak, Kretch, Soska, & Adolph, 2011; Frank, 2012; Frank, Simmons, Yurovsky, & Pusiol, 2013; Smith, Yu, & Pereira, 2011; Yoshida & Smith, 2008). Recent studies using these methods suggest that the visual world of older infants may be more complex than that of younger infants, especially when it comes to finding the faces of people around them.

One example of the increasing complexity of infants' visual world comes from Aslin (2009), who showed groups of 4- and 8-month-olds videos collected from a head-mounted camera worn by a child of approximately the same age. Whereas the 4-month-olds looked at roughly the same locations in the age-appropriate videos as a control group of adults did, the 8-month-olds looked at the faces of people in the videos significantly less than adults did. Perhaps in supportive contexts where caregivers routinely hold infants close to their own faces, there is a good match between the more limited attentional abilities of young infants and their more restricted visual environment (Stern, 1977). In contrast, in contexts of early neglect or deprivation, the mismatch between children's limited attentional abilities and the relative distance of the faces of caregivers might lead to pervasive long-term difficulties in face processing (e.g., as in children raised in institutions; Moulson, Westerlund, Fox, Zeanah, & Nelson, 2009). Although it seems unlikely that visual stimulation alone—in the absence of a supportive family environment—would remediate such difficulties, our argument here is simply that appropriate visual stimulation is an integral part of such supportive environments.

Our work adds to a growing body of evidence suggesting that even perception requires practice. Knowledge that is manifest through longer looking times may be mediated by the ability to attend systematically to the appropriate aspects of the stimulus. For example, infants are more likely to perceptually complete an object behind an occluder when they are more accurate at visual search and when that in turn leads them to scan back and forth between halves of the occluded object (Amso & Johnson, 2006; Johnson, Slemmer, & Amso, 2004; Johnson et al., 2008). We have argued here that a similar relation may hold during the early development of social attention; being able to learn from social signals requires the attentional wherewithal to find them.

Acknowledgments

This work was supported by National Institutes of Health (NIH) Grants R01-HD40432 and R01-HD73535 to S.P.J., supporting data collection, and a grant from the John Merck Scholars Fund to M.C.F., supporting data analysis. Thanks go to Ed Vul for valuable discussion, to Rahman Zolfaghari for assistance with data collection, and to the infants and families who participated in the study.

References

Amso, D., & Johnson, S. (2006). Learning by selection: Visual search and object perception in young infants. *Developmental Psychology, 42*, 1236–1245.

Amso, D., & Johnson, S. (2008). Development of visual selection in 3- to 9-month-olds: Evidence from saccades to previously ignored locations. *Infancy, 13*, 675–686.

Aslin, R. (2009). How infants view natural scenes gathered from a head-mounted camera. *Optometry & Vision Science, 86*, 561–565.

Bahrick, L., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology, 36*, 190–201.

Bahrick, L., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Current Directions in Psychological Science, 13*, 99–102.

Baron, J. (1979). Orthographic and word-specific mechanisms in children's reading of words. *Child Development, 50*, 60–72.

Baron, R., & Kenny, D. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.

Cashon, C. H., Ha, O.-R., Allen, C. L., & Barna, A. C. (2013). A U-shaped relation between sitting ability and upright face processing in infants. *Child Development, 84*, 802–809.

Cicchino, J., Aslin, R., & Rakison, D. (2011). Correspondences between what infants see and know about causal and self-propelled motion. *Cognition, 118*, 171–192.

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology, 52*, 337–367.

Dannemiller, J. (1998). A competition model of exogenous orienting in 3.5-month-old infants. *Journal of Experimental Child Psychology, 68*, 169–201.

Dannemiller, J. (2000). Competition in early exogenous orienting between 7 and 21 weeks. *Journal of Experimental Child Psychology, 76*, 253–274.

Dannemiller, J. (2005). Motion popout in selective visual orienting at 4.5 but not at 2 months in human infants. *Infancy, 8*, 201–216.

DeNicola, C. A., Holt, N. A., Lambert, A. J., & Cashon, C. H. (2013). Attention-orienting and attention-holding effects of faces on 4- to 8-month-old infants. *International Journal of Behavioral Development, 37*, 143–147.

Di Giorgio, E., Turati, C., Altoè, G., & Simion, F. (2012). Face detection in complex visual displays: An eye-tracking study with 3- and 6-month-old infants and adults. *Journal of Experimental Child Psychology, 113*, 66–77.

Escudero, P., Robbins, R. A., & Johnson, S. P. (2013). Sex-related preferences for real and doll's faces versus real and toy objects in young infants and adults. *Journal of Experimental Child Psychology, 116*, 367–379.

Farroni, T., Johnson, M., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005). Newborns' preference for face-relevant stimuli: Effects of contrast polarity. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 17245–17250.

Franchak, J., Kretch, K., Soska, K., & Adolph, K. (2011). Head-mounted eye tracking: A new method to describe infant looking. *Child Development, 82*, 1738–1750.

Frank, M. C. (2012). Measuring children's visual access to social information using face detection. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th annual meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Frank, M. C., Simmons, K., Yurovsky, D., & Pusiol, G. (2013). Developmental and postural changes in children's visual access to faces. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual meeting of the Cognitive Science Society*. Cognitive Science Society: Austin, TX.

Frank, M. C., Vul, E., & Johnson, S. (2009). Development of infants' attention to faces during the first year. *Cognition, 110*, 160–170.

Frank, M. C., Vul, E., & Saxe, R. (2012). Measuring the development of social attention using free-viewing. *Infancy, 17*, 355–375.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naive theory of rational action. *Trends in Cognitive Sciences, 7*, 287–292.

Gliga, T., Elsabbagh, M., Andravizou, A., & Johnson, M. (2009). Faces attract infants' attention in complex displays. *Infancy, 14*, 550–562.

Gluckman, M., & Johnson, S. P. (2013). Attentional capture by social stimuli in young infants. *Frontiers in Psychology, 4*, 527.

Johnson, S. P., Amso, D., & Slemmer, J. (2003). Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 10568–10573.

Johnson, S. P., Davidow, J., Hall-Haro, C., & Frank, M. (2008). Development of perceptual completion originates in information acquisition. *Developmental Psychology, 44*, 1214–1224.

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.

Johnson, S. C., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science, 1*, 233–238.

Johnson, S. P., Slemmer, J., & Amso, D. (2004). Where infants look determines how they see: Eye movements and object perception performance in 3-month-olds. *Infancy, 6*, 185–201.

Kelly, D., Quinn, P., Slater, A., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science, 18*, 1084–1089.

Libertus, K., & Needham, A. (2011). Reaching experience increases face preference in 3-month-old infants. *Developmental Science, 14*, 1355–1364.

Macchi, C. V., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science, 15*, 379–383.

Moulson, M., Westerlund, A., Fox, N., Zeanah, C., & Nelson, C. (2009). The effects of early experience on face recognition: An event-related potential study of institutionalized children in Romania. *Child Development, 80*, 1039–1056.

Pascalis, O., de Haan, M., & Nelson, C. (2002). Is face processing species-specific during the first year of life? *Science, 296*, 1321–1323.

Pascalis, O., Scott, L., Kelly, D., Shannon, R., Nicholson, E., Coleman, M., et al (2005). Plasticity of face processing in infancy. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 5297–5300.
Richards, J. (2010). The development of attention to simple and complex visual stimuli in infants: Behavioral and psychophysiological measures. *Developmental Review, 30*, 203–219.
Richards, J., & Holley, F. (1999). Infant attention and the development of smooth pursuit tracking. *Developmental Psychology, 35*, 856–867.
Saxe, R., Tenenbaum, J., & Carey, S. (2005). Secret agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science, 16*, 995–1001.
Scaife, M., & Bruner, J. (1975). The capacity for joint visual attention in the infant. *Nature, 253*, 265–266.
Simion, F., Cassia, V., Turati, C., & Valenza, E. (2001). The origins of face perception: Specific versus non-specific mechanisms. *Infant and Child Development, 10*, 59–65.
Smith, L., Yu, C., & Pereira, A. (2011). Not your mother's view: The dynamics of toddler visual experience. *Developmental Science, 14*, 9–17.
Stern, D. (1977). *The first relationship: Infant and mother*. Cambridge, MA: Harvard University Press.
Treiman, R. (1984). Individual differences among children in spelling and reading styles. *Journal of Experimental Child Psychology, 37*, 463–477.
Triesch, J., Teuscher, C., Deák, G., & Carlson, E. (2006). Gaze following: Why (not) learn it? *Developmental Science, 9*, 125–147.
Turati, C., Valenza, E., Leo, I., & Simion, F. (2005). Three-month-olds' visual preference for faces and its underlying visual processing mechanisms. *Journal of Experimental Child Psychology, 90*, 255–273.
Walden, T., & Ogan, T. (1988). The development of social referencing. *Child Development, 59*, 1230–1240.
Yoshida, H., & Smith, L. (2008). What's in view for toddlers? Using a head camera to study visual experience. *Infancy, 13*, 229–248.