Infants' responses to interactive gaze-contingent faces in a novel and naturalistic eye-tracking paradigm

Authors

Jolie. R. Keemink, Maryam Keshavarzi-Pour and David J. Kelly

Author affiliations

School of Psychology, Keynes College, University of Kent, Canterbury, Kent, CT2 7NP

Corresponding Author

Jolie. R. Keemink, School of Psychology, Keynes College, University of Kent, Canterbury,

Kent, CT2 7NP

Email: jrk26@kent.ac.uk

Disclosures

All authors report having no potential conflict of interest.

Abstract

Background: Face-scanning is an important skill that takes place in a highly interactive context embedded within social interaction. However, previous research has studied face-scanning using non-interactive stimuli. We aimed to study face-scanning and social interaction in infancy in a more ecologically valid way by providing infants with a naturalistic and socially engaging experience.

Methods: We developed a novel gaze-contingent eye-tracking paradigm in which infants could interact with face-stimuli. Responses (socially engaging/socially disengaging) from faces were contingent on infants' eye movements. We collected eye-tracking and behavioral data of 162 (79 male, 83 female) 6-, 9- and 12-month-olds.

Results: All infants showed a clear preference for looking at the eyes relative to the mouth. Contingency was learned implicitly and infants were more likely to show behavioral responses (e.g. smiling, pointing) when receiving socially engaging responses. Infants' responses were also more often congruent with the actors' responses than incongruent. Additionally, our large sample allowed us to look at the ranges of behavior on our task and we identified a small number of infants who displayed deviant behaviors. We discuss these findings in relation to data collected from a small sample (N=11) of infants considered to be 'at-risk' for autism spectrum disorders.

Conclusions: Our results demonstrate the versatility of the gaze-contingency eye-tracking paradigm, allowing for a more nuanced and complex investigation of face-scanning as it happens in real-life interaction. As we provide additional measures of contingency learning and reciprocity, our task holds the potential to investigate atypical neurodevelopment within the first year of life.

Keywords: Face scanning, gaze-contingent, eye-tracking, autism spectrum disorder, infant siblings, early detection

Face scanning in infancy

Faces represent a stimulus category of unique importance generating greater attention compared to other visual stimuli (Kelly et al., under review; Langton, Law, Burton & Schweinberger, 2008). Newborn human infants show a preference for faces and face-like stimuli (Johnson, Dziurawiec, Ellis & Morton, 1991; Mondloch et al., 1999; Valenza, Simion, Macchi Cassia & Umiltà, 1996), recognise and prefer their mother’s face (Bushnell, Sai & Mullin, 1989; Pascalis, De Schonen, Morton, Deruelle & Fabre-Grenet, 1995), and favor attractive faces (Slater et al., 1998). Infants show particular interest in the eye region (Di Giorgio, Méary, Pascalis & Simion, 2013; Haith, Bergman & Moore, 1977; Maurer & Salapatek, 1976) and from birth engage in and actively search for mutual eye gaze (Farroni, Csibra, Simion, & Johnson, 2002). There is a rapid increase in attention to faces between 3 and 11 weeks of age (Haith et al., 1977) with an eye preference present in 6-week-old infants (Hunnius & Geuze, 2004). Several studies report a subsequent decrease in eye region attention from 6-12 months of age (Lewkowicz & Hansen-Tift, 2011; Tenenbaum, Shah, Sobel, Malle & Morgan, 2013) with infants shifting their focus to the mouth region, attributed to language learning. Similarly, Oakes and Ellis (2013) demonstrated an eye preference in 4.5-6.5-months-old infants and more distributed looking in older 8- to 12-months-old infants. Collectively, these studies have provided insights into face scanning throughout the first year of life, yet the extent to which their methodologies produce findings that generalize to ‘real world’ social interactions is unclear.

Methodologies

When infants encounter faces outside of the lab, this takes place in a highly interactive, social context in which reciprocity and contingency play a crucial role. However, past studies have attempted to answer questions about face-to-face *interactions*, whilst using methods that employ *non-interactive* stimuli. Although previous methods present infants with facial stimuli, such as static images (Di Giorgio et al., 2013; Oakes and Ellis, 2013), videos of faces (Hunnius and Geuze, 2004; Lewkowicz & Hansen-Tift, 2011; Tenenbaum et al., 2013) and real faces (Haith et al., 1977; Maurer & Salapatek, 1976), these stimuli do not capture the reciprocity inherent to the social context in which face scanning occurs. By reducing face scanning to an isolated skill, we lose the richness and meaningfulness of the interactive context. In order to overcome this methodological issue, the current study introduces a novel eye-tracking paradigm in which infants are presented with interactive, gaze-contingent (GC) faces whilst their behavioural responses (e.g. smiles, head shaking) towards the interactive faces are measured.

The Gaze-contingency paradigm

Advances in eye-tracking permit the fine-grained study of infants' responses to visual stimuli and enable the implementation of novel and interactive GC paradigms. In GC paradigms the participant's viewing experience is contingent upon their eye movements, which allows the participant to 'interact' with stimuli providing a more naturalistic and interactive experience. A small number of studies have indicated that GC paradigms can be effectively implemented in adult and infant research (Deligianni, Senju, Gergely & Csibra, 2011; Miyazaki, Takahashi, Rolf, Okada & Omori, 2014; Wang et al., 2012; Wilms et al., 2010). Furthermore, previous research has established that from 2 months of age, infants are sensitive to and are capable of learning visual (De Schonen & Bry, 1987; Johnson, Posner & Rothbart, 1991), social (Rochat, Querido & Striano, 1999; Soussignan, Nadel, Canet & Gerardin, 2006) and physical contingencies (Alessandri, Sullivan & Lewis, 1990; Angulo-Kinzler, Ulrich & Thelen, 2002; Rovee & Rovee, 1969). Face scanning lends itself perfectly for GC paradigms because of its interactive nature. However, surprisingly, there are no published studies to date investigating face scanning using GC paradigms.

The current study comprises a novel and unique combination of gaze-contingent eye-tracking and behavioral measures designed to capture social interaction in a controlled lab environment and to establish the efficacy of this paradigm within infant face scanning research. The task will simultaneously provide measures of initial fixation location, contingency learning, face scanning and behavioral responses (i.e. reciprocity). Furthermore, by testing a large sample of typically developing infants, ranges of typical behavior will be established. In addition to contrasting groups, a distribution-based approach allows us explore and establish the ranges of typical behaviors within a GC paradigm. Contrary to previous studies that have explored face processing strategies in only one age group (Young, Merin, Rogers, & Ozonoff, 2009), and with limited sample sizes (e.g., Klin et al., 2002), the current study sample comprised 6-, 9-, and 12-month-olds (*n* = 162).

For the task, participants sequentially viewed a series of video-recorded actors that could produce either a socially engaging or a socially disengaging response, which is contingent on first fixation location. Based on existing face scanning research (e.g., Di Giorgio et al., 2013), it was hypothesized that infants would be likely to initially fixate the eye region. However, previous findings (e.g., Soussignan et al., 2006) also led us to hypothesize that infants might be capable of learning the task contingency and consequently would favor triggering socially engaging responses. Finally, infants' faces were video-recorded throughout testing. Infants were expected to show behavior congruent with the triggered response from the actors (e.g., a smile for a smile) (Hains & Muir, 1996).

Extension

The novel, interactive nature of our GC stimuli produces a more socially demanding task compared to previous methods, which lead us to believe that eventually the task could be employed to explore early signs of atypical social development. A recent line of research has focused on infants at high familial risk for ASD (HR; because of an older sibling with a diagnosis), allowing for prospective investigation of the development of ASD. Several studies looking at face scanning in HR infants suggest some deviancies (Chawarska, Macari, Shic, 2013; Guiraud et al., 2012; Shic, Macari & Chawarska, 2014; Merin, Young, Ozonoff & Rogers, 2007), although Young et al. (2009) report that these are not related to later ASD outcomes. However, these studies employ *non-interactive* stimuli similar to the aforementioned research. In contrast to any previously published research, our design will enable us to explore discrete social interactions within a controlled laboratory setting by synthesizing fine-grained eye movement analyses with overt behavioral reactions, permitting more meaningful conclusions about face scanning in typical and atypical populations. Therefore, for exploratory purposes a small sample of HR infants will be included and compared to the established behavioral norms for the typical population.

Methods and Materials

This study was approved by the Ethics Committee of the University of Kent (Protocol number: 20153600, Project name: Social Interaction preferences and visual face scanning strategies in 6-12-months-olds: evidence from a gaze-contingency paradigm). All parents signed an informed consent for their participating infant. Data were stored and treated anonymously.

Participants

Typically developing infants were recruited through the Kent Child Development Unit database of families who have enlisted for research. Infants were considered typically developing if they had no known medical/psychological conditions. The final sample consisted of 162 infants (79 male, 83 female), who were separated into three different age groups: 6-month-olds, 9-month-olds and 12-month-olds (See Table 1). All infants were Caucasian. Infants were randomly assigned to either the *Social Eyes* (*SE*) condition (*n* = 89) or the *Social Mouth* (*SM*) condition (*n* = 73). A further 16 infants participated in the SM condition, but disengaged from the task. Eleven infants in our sample were classified as HR, as they had an older sibling with a formal diagnosis. They were recruited through autism support groups across Kent.

After group analyses were conducted on the total sample, we explored individual performance of HR infants. Previous research looking at early markers for ASD has contrasted a HR sample with a TD sample to examine group differences. In addition to investigating group differences, we propose it might be meaningful to examine individual HR behavior compared to ranges of typical behavior. Arguably, HR infants do not constitute a separate group (yet), as only ~20% of them will receive an ASD diagnosis (Ozonoff et al., 2011). It is that subgroup that potentially will differentiate from the typical range on sufficiently sensitive measures.

The Gaze-Contingent Task

The GC task consisted of a series of video-recorded actors who could produce a response of low or high social engagement (closed/open smile), or a response of low or high social *dis*engagement (closing eyes/looking away) contingent on the infant's first fixation location. The behavior produced by the actors was contingent on the infant's eye movements and triggered by engaging in eye contact or fixating on their mouth (See Figure 1). The animation was triggered from the first fixation landing in either of these regions. We chose to include both socially engaging and disengaging responses to explore whether infants were motivated to seek out a socially engaging response, and to investigate a potential difference in behavioral responses. In the *SE* condition, infants triggered socially engaging responses by fixating the eyes and socially disengaging responses by fixating on the mouth. Responses were reversed in the *SM* condition (See Figure 2). If an infant did not fixate on the discrete regions within the trial-length, the face would not animate.

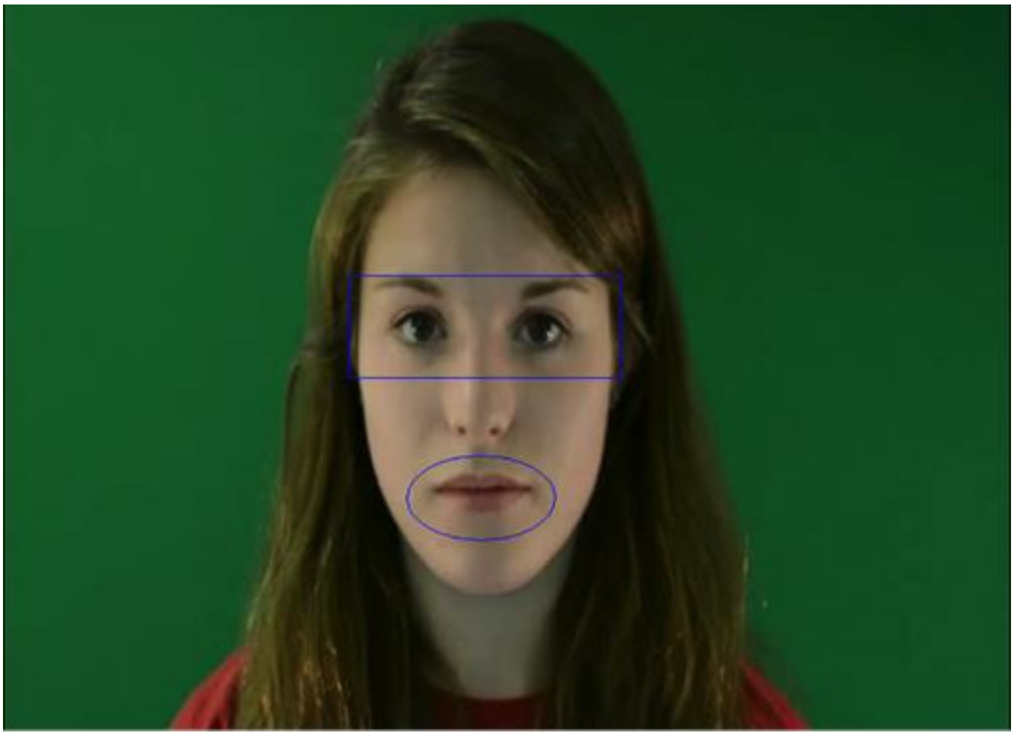
Table 1

Participant Characteristics per Age group and Condition

	Age in Months	*Condition*	*N*	*Mean Age in Days (SD)*	*Age Range*	*Gender (M/F)*
TD	6	SE	29	198 (4.8)	187 - 206	(12/17)
		SM	21	199 (4.5)	188 - 209	(10/11)
	9	SE	29	279 (9.1)	263 - 303	(13/16)
		SM	24	280 (7.7)	266 - 293	(13/11)
	12	SE	28	370 (9.4)	354 - 387	(14/14)
		SM	20	371 (9.1)	354 - 388	(8/12)
HR	6	SE	-	-	-	-
		SM	3	190 (2.1)	188 – 192	(3/0)
	9	SE	1	283 (N/A)	N/A	(1/0)
		SM	-	-	-	-
	12	SE	2	367 (10.6)	360 - 375	(2/0)
		SM	5	370 (5.6)	362 - 377	(3/2)
Total		SE	89			(42/47)
		SM	73			(37/36)

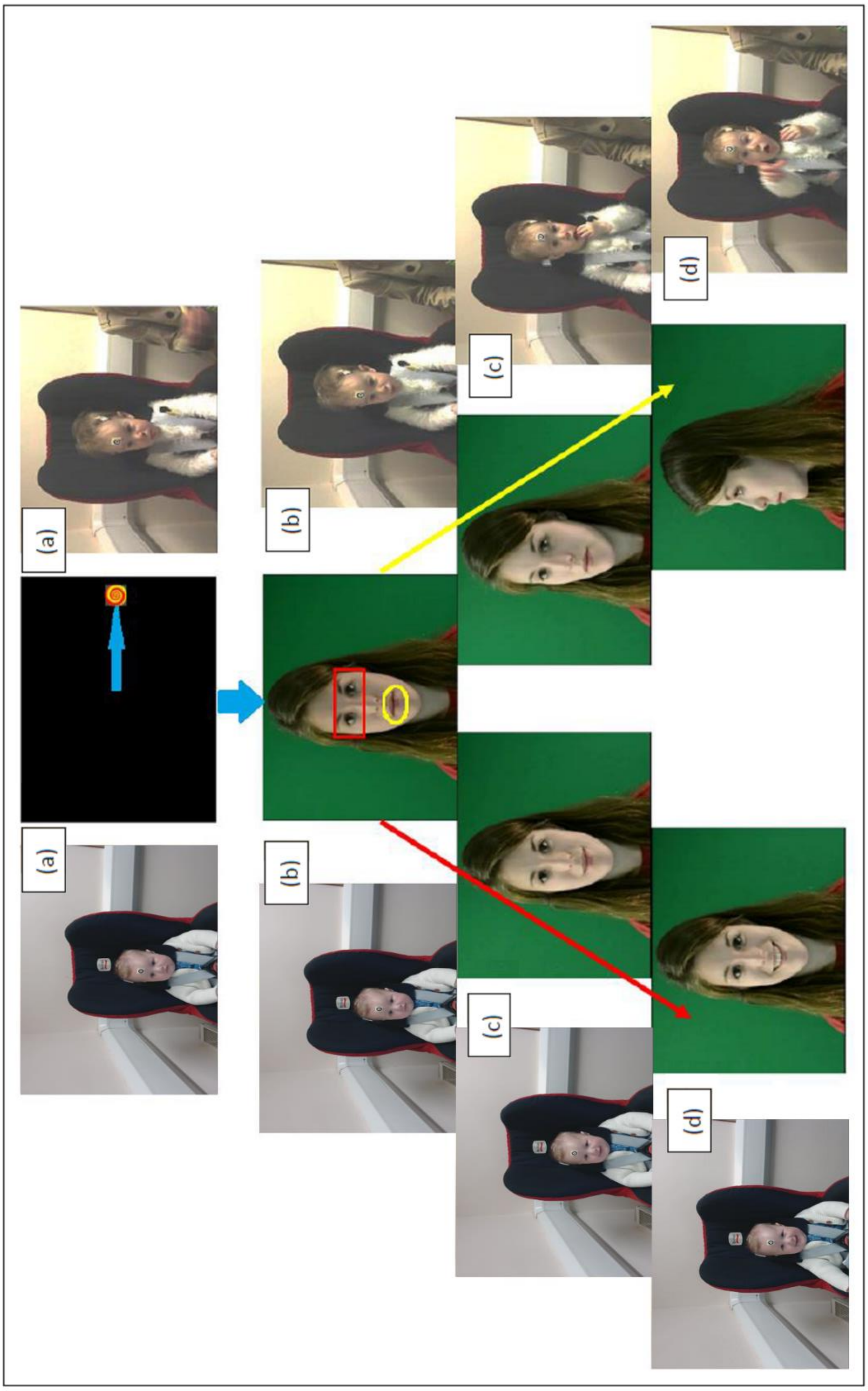
Description of Stimuli

The stimuli were 20 color videos of ten neutral-looking male and female adult faces visible from the shoulders upward standing in front of a green screen (See Figure 1). Each stimulus appeared twice and in consecutive trials to assess if learning occurred across presentations. Each trial lasted five seconds. Eight faces were of Caucasian origin and two faces of African origin. All images subtended a size of 24.77 degrees x 18.25 degrees in visual angle and were presented on a 20-inch monitor with a resolution of 1024 by 768 pixels. Discrete gaze-contingent 'invisible boundaries' for eye and mouth regions were defined individually for each face (See Figure 1). All eye regions measured 6.8 x 2.83 degrees and all mouth regions measured 5.06 x 2.83 degrees. A dissimilarity in AOI size is common practice in



infant face scanning research (e.g. Chawarska et al., 2013; Wagner, Luyster, Tager-Flusberg & Nelson, 2016).

Figure 1. A Stimulus with the Discrete Eye and Mouth Regions Visible. The individual whose face appears here gave signed consent for her likeness to be published in this article.



(a). The infant fixates a target that appears randomly to the left or right of the screen, which triggers the appearance of a face.
(b). A fixation landing in one of the two Areas of Interest (AOIs: boundaries are NOT visible during test) causes the face to immediately animate. In each trial, infants could trigger either a socially-engaging (e.g., smile) or a socially non-engaging (e.g., head turning) animation. The animation type triggered by eye and mouth AOIs was counterbalanced across infants (i.e., eye fixations produced smiles for half the infants and head-turns for the other infants; vice versa for the mouth).
(c – d). The animation occurs and the infant reacts. The stills captured from video recordings provide examples of congruous infant responses to social and non-social stimulus animations.

Figure 2. Examples of Socially-Engaging and Socially-Disengaging Animations with Accompanying Infants' Behavioral Responses. The authors received signed consent for the woman's and children's likenesses to be published in this article.

Equipment

Eye movements were recorded with an Eyelink 1000+ (SR Research, Ontario) at a sampling rate of 500 Hz operated in Remote Mode using a 25mm lens attachment. Infants aged 12 months were tested using the 890 nm illuminator, while all other age groups were tested using the 940 nm illuminator. Under optimal conditions, when operating in Remote Mode the Eyelink has accuracy of 0.5°, a tracking range of 32° (horizontal) x 25° (vertical) and is tolerant to head movements of 22x18x20cm. In order to minimise head movements, infants were securely fastened in an age-appropriate car seat that was safely attached to a chair. Stimuli were presented using Experiment Builder (SR Research, Ontario) and the raw eye movement data were extracted using Data Viewer (SR Research, Ontario). Fixations and saccades were subsequently parsed in Matlab (The Mathworks, MA, USA) using custom written code (See Supplementary Information for further details). All subsequent data processing was conducted in Matlab.

In addition, infants' behavioral responses were recorded with a Logitech webcam. Recordings were analyzed frame-by-frame and coded by one of the researchers and an independent coder (see below). Agreement between the coders was .94 (See Supplementary Information for additional information).

Procedure

Families were welcomed and informed about the study. Parents were asked to sign a consent form and then escorted to the research laboratory with dimmed lighting. Infants were placed in an age-appropriate seat at a viewing distance of 60cm from a computer monitor. The infant's right eye was tracked throughout testing. Infants' behavioral responses were also recorded throughout. The infant’s view to their surroundings and experimenters was obstructed by an occluding screen. A 5-point calibration procedure using custom-made attention-grabbing audio-visual targets (Supplied by Dr. David Meary) was conducted and repeated as necessary. To ensure that all eye movement data was accurate, all infants were calibrated and validated to within 1° and checks for drift were assessed between every single trial. No infant failed to calibrate. Following calibration, the task was initiated. An attention grabber appeared at the side of the screen between each stimulus presentation that ensured the infant’s gaze for the beginning of each trial. The study lasted approximately five minutes. Infants received a young scientist certificate and a small age-appropriate gift.

Behavioral Coding

Video-recorded behavioral responses could be categorized as positive, negative, ambiguous or a non-response. Smiling, waving, giggling, cheerful vocalizing and cheerful pointing were seen as positive responses. Negative responses comprised looking away, vocalizing, frowning, head shaking and sad facial expressions with some of the older infants showing more complex behaviors such as indignant pointing. Some responses fell in-between categories and were coded as ambiguous (e.g. arbitrary head movements). If the infant maintained a neutral facial expression throughout the trial, the trial was coded as ‘no response’.

Subsequently, eye movement data were time-locked with the behavioral data to ensure the infant’s behavior occurred in *response* to the triggered animation and whether it was congruent or incongruent with the triggered animation. Congruent responses comprised a positive behavior from the infant towards an actor’s socially engaging response or a negative behavior from the infant towards an actor’s socially disengaging response. Conversely, incongruent responses were a positive behaviors towards a socially disengaging response or a negative behavior from the infant towards a socially engaging response (See Figure 2 for examples of congruent responses).

Results

As aspects of our study methodology are completely novel, we were unable to conduct accurate *a priori* power analyses, but *post-hoc* power analyses indicated very high power (.88 - .99) for all main effects and interactions with the exception of Condition x Response for behavioral responses, which was notably low (.15). Preliminary analyses indicated no differences of participant gender, so it was omitted for further analyses. Ethnicity and stimulus gender did not affect any infant responses, nor did face repetition. Eye movement analyses will first be described covering overall AOI dwell time followed by explicit (i.e. the percentage of fixations triggering socially engaging responses) and implicit (i.e. saccadic response times across trials) measures of contingency learning. Behavioral responses will subsequently be assessed. Data will be analyzed with 3 (Age: 6, 9 or 12 months) x 2 (Condition: SE or SM) x 2 (AOI: Eyes or Mouth) ANOVAs and appropriate post-hoc tests unless stated otherwise. Finally, Z-normalized ranges of typical behavior on key dependent measures will be assessed and the performance of a small sample of HR infants will be contrasted typical norms.

Eye-movement data

Overall AOI Dwell Time

The independent measures for these analyses are age, condition and AOI. The dependent measure is overall percentage of fixations.

We analyzed whether dwell time to eyes and mouth (cumulative time following animation trigger) differed by condition and/or age. A univariate ANOVA revealed a significant effect of *Condition* ($F(1,312) = 7.100$, $p = .008$, $\eta_p^2 = .022$), a main effect of *AOI* ($F(1,312) = 389.828$, $p < .001$, $\eta_p^2 = .555$) and a significant *Condition x AOI* interaction ($F(1,312) = 9.105$, $p = .003$, $\eta_p^2 = .028$). Inspection of means confirmed that eyes were fixated more (SE = 39.24%; SM = 39.71%) than the mouth (SE = 17.86%; SM = 10.57%) by infants of all ages and regardless of condition (See Additional Analyses 1 in Supplementary Information).

Independent t-tests confirmed that dwell time did not differ between conditions for Eyes ($t(160) = -.2588$, $p = .797$). However, the mouth was fixated significantly more ($t(160) = 3.889$, $p <.001$) by infants in the *SE* condition than the *SM* condition. Although seemingly counterintuitive, this finding can be accounted for by the tendency for infants to look directly at the mouth once a smile was initiated (See Figure 3).

Separate analyses for dwell time on eyes and mouth yielded a significant effect of age on dwell time on the eye area ($F(2,159) = 4.080$, $p = .019$, $\eta_p^2 = .049$), but not for the mouth area ($F(2,159) = 1.303$, $p = .275$, $\eta_p^2 = .016$). Post-hoc analyses revealed differences between 6- and 12-month-olds only ($p = .005$), with 6-month-olds fixating the eye area more ($M = 42.25\%$) relative to 12-month-olds ($p = 36.20\%$).

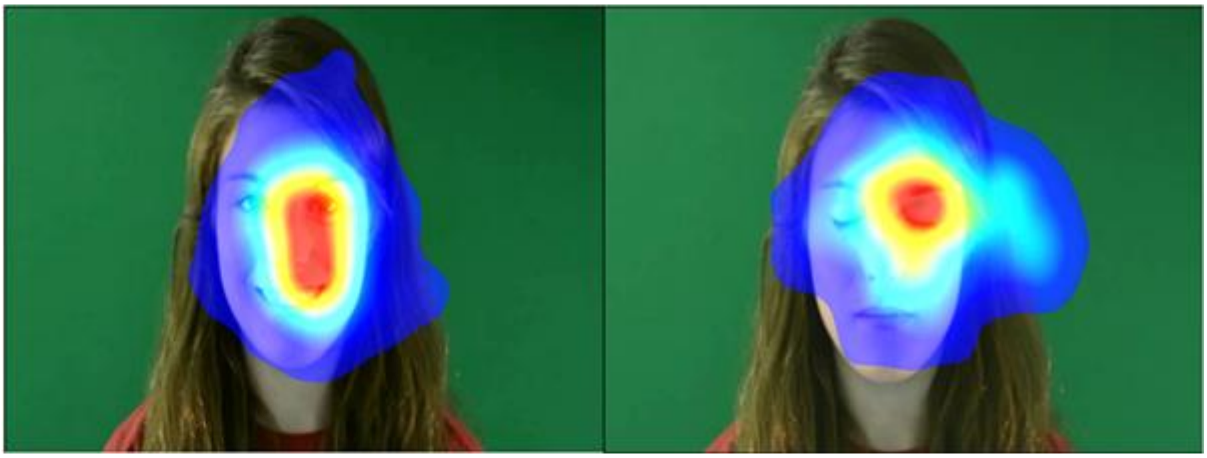


Figure 3. Summed fixations post-animation onset for all infants in the *SE* condition on the left and *SM* condition on the right. The tendency for infants to fixate the mouth (smile) in the *SE* condition but not in the *SM* condition is clearly visible. The individual whose face appears here gave signed consent for her likeness to be published in this article.

Explicit Contingency Learning: First Fixation Location

Explicit contingency learning would be demonstrated if across trials infants' first fixations more frequently fell within the area that produced a socially engaging response (eyes for the *SE* condition, mouth for the *SM* condition). This would indicate that infants had learned what area to fixate in order to trigger social engagement. The independent measures for these analyses are age, condition and AOI (mouth vs eyes). The dependent measures are the percentages of first fixations.

A univariate ANONA revealed a main effect of AOI ($F(1,312) = 533.842$, $p < .001$, $\eta_p^2 = .631$) and a significant Age x AOI interaction ($F(2,312) = 6.567$, $p = .002$, $\eta_p^2 = .040$).

Regardless of condition and age, infants were far more likely to initially fixate the eyes relative to the mouth (*SE* eyes: 71.48%, mouth: 18.28% and *SM* eyes: 69.21%, mouth: 15.41%). Post-hoc one-way ANOVAs revealed significant Age differences for Eyes only, with significant differences between 6- and 12-month olds (p = .003); 12-month olds displayed fewer eye fixations (M = 62.20) relative to 6-month-olds (M = 75.33). However, regardless of age or condition, infants showed a clear tendency to initially fixate the eyes relative to the mouth. Evidence for explicit contingency learning was not found.

Implicit Contingency Learning: Saccadic Latencies

In addition to explicit contingency learning, we investigated implicit contingency learning. We analyzed saccadic response times to trigger the face animations. We reasoned that, as an implicit response from the infant to the two different social responses of the actors, infants could become more eager (faster saccades) or more reluctant (slower saccades) to trigger the animations. Infants in the *SE* condition would demonstrate implicit contingency learning if their saccadic response times in trials 11-20 were *faster* relative to trials 1-10 as a result of the socially engaging responses. Infants in the *SM* condition demonstrated implicit contingency learning if their saccadic response times in trials 11-20 were *slower* relative to trials 1-10 because of the socially disengaging responses.

Saccadic response times to trigger the animation were calculated and the saccadic response times for trials 1-10 and trials 11-20 were contrasted, as previous studies have indicated that infants show evidence of learning within 10 trials (e.g., Colombo, Mitchell, Coldren & Atwater, 1990; Fawcett & Liszkowski, 2012; Hauf & Aschersleben, 2008). The independent measures for these analyses are trials (1-10 vs. 11-20) and condition. The dependent measure is the saccadic response time (in seconds) it took infants to trigger the animation. (N.B. Animations could be triggered by fixating either the eye or the mouth area, so the DV comprises saccades to either of these areas. In reality, saccades more often went to the eye area as demonstrated in the section on first fixation location).

Preliminary analysis demonstrated no overall significant differences in saccadic response time between conditions (*F*(1,150) = 1.056, *p* = .795, $\eta_p^2 < .001$). Subsequently, a 2 (Trials) x 2 (Condition) repeated measures ANOVA was conducted on the saccadic response times split across trials 1-10 and trials 11-20. The ANOVA yielded a significant Trials x Condition interaction (*F*(1,156) = 9.724, *p* = .002, $\eta_p^2 = .059$). Inspection of means revealed that saccadic response times did not differ between conditions for trials 1-10 (*SE* = 693 msecs, *SM* = 651 msecs), but they differed substantially for trials 11-20 (*SE* = 605 msecs, *SM* = 712 msecs).

Further two-tailed *t*-tests confirmed implicit contingency learning as summarized in Figure 4. Independent-samples t-tests analyzing differences in response times revealed no difference between conditions for trials 1-10 (*t*(160) = .938, *p* = .35), whereas for trials 11-20 infants in the *SE* condition demonstrated significantly faster saccades compared to infants in the *SM* condition as a result of socially engaging and socially disengaging responses respectively (*t*(160) = -2.660, *p* = .009). Additionally, paired-samples t-tests analyzing differences between trials 1-10 and trials 11-20 within conditions, revealed a significant difference for the *SE* condition only (*t*(88) = 2.711, *p* = .008). Infants in the *SE* condition showed significant faster saccades on trials 11-20 compared to trials 1-10, whereas there was no difference for the *SM* condition (*t*(72) = -1.690, *p* = .094).

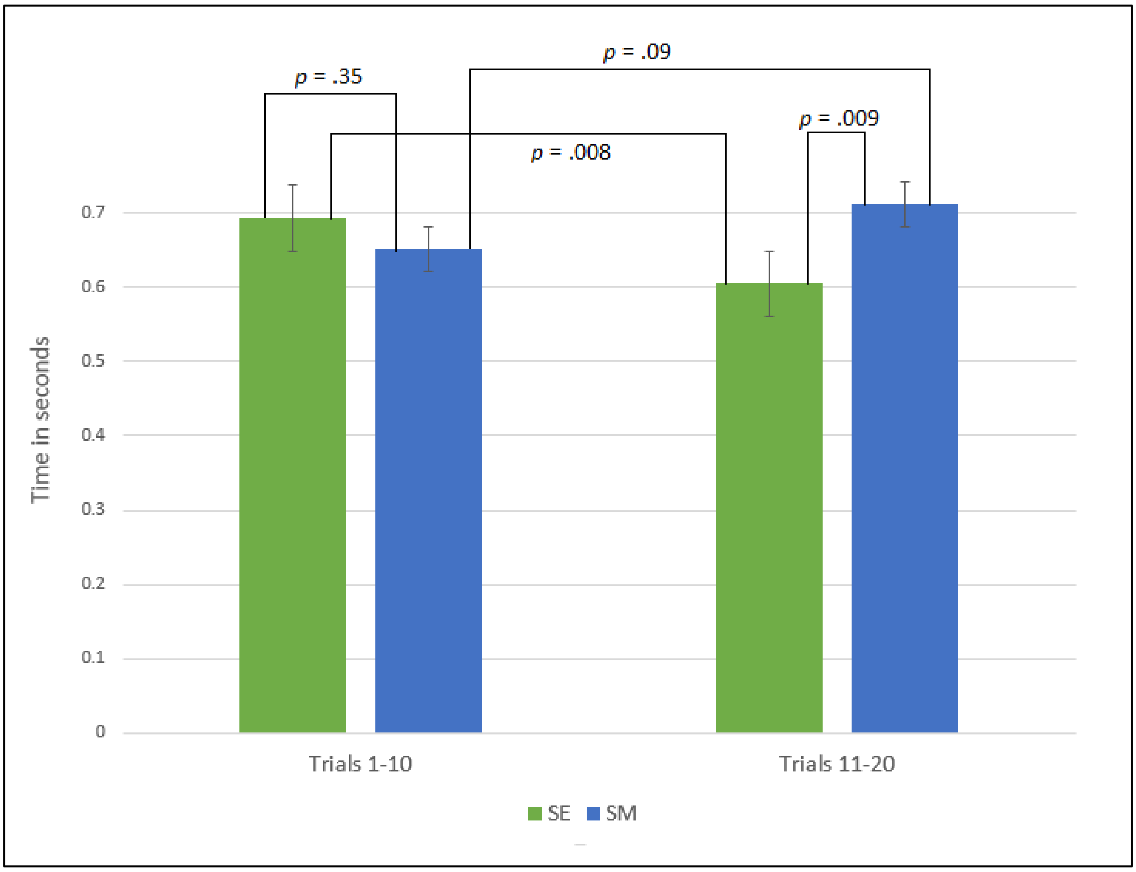


Figure 4. Average Saccadic Response Times for Both Conditions on Trials 1-10 and Trials 11-20.

Between conditions there was no difference in response times for trials 1-10, but there was a significant difference for trials 11-20, implicating implicit contingency learning. Within conditions, there was a significant difference in saccadic response times between trials 1-10 and trials 11-20 for only the *SE* condition.

Behavioral responses

To assess behavioral responses we determined if infants' behavior was congruent or incongruent with the actor's triggered response. As a consequence of recording errors, the behavioral data from six infants (2 x 6m, 3 X 9m & 1 x 12m) was lost. Preliminary analysis of the remaining data (SE: $n = 83$; SM: $n = 73$) revealed no effects of gender, so data were collapsed for further analyses (See Supplementary Table 1). As data were highly skewed, a log transform was conducted prior to performing data analyses.

A 3 (Age) x 2 (Condition) x 2 (Response Type; Congruent, Incongruent) univariate ANOVA conducted on percentage of responses revealed a main effect of *Condition* ($F(1,300) = 18.869$, $p < .001$, $\eta_p^2 = .059$) and *Response Type* ($F(1,300) = 91.239$, $p < .001$, $\eta_p^2 = .233$), a significant *Age x Condition* interaction ($F(2,300) = 10.579$, $p < .001$, $\eta_p^2 = .066$) and a *Condition x Response Type* interaction ($F(2,300) = 11.574$, $p < .001$, $\eta_p^2 = .037$). Inspection of means shows that Congruent responses ($M = 16.00\%$) were observed more frequently than Incongruent responses ($M = 10.21\%$) and that the infants were more likely to respond in the *SE* condition ($M = 15.47\%$) relative to the *SM* condition ($M = 10.27\%$). In terms of age related differences, post-hoc comparisons found that only 9-month-olds ($M = 15.12\%$, $p = .036$) responded more frequently relative to 6-month-olds ($M = 10.40\%$).

To explore the interactions, separate univariate ANOVAs were conducted for the *SE* and *SM* condition, which yielded age-related differences in the *SE* condition only ($F(2,167) = 6.399$, $p = .002$, $\eta_p^2 = .071$). Post hoc comparisons found that 6-month olds were less likely to respond ($M = 8.81\%$) relative to both 9-month-olds ($M = 20.64\%$; $p = .001$) and 12-month-olds ($M = 16.51\%$; $p < .023$).

Distribution of Performance and HR Comparison

Having tested a large sample of typically developing infants in a novel research paradigm, we decided to establish ranges for typical behavior. To assess behavior of individual HR infants we produced z-normalized scores and distributions for key measures. Constructing z-normalized distributions, it is notable that different measures produced different distribution shapes; normal and skewed. A normal distribution indicates that a behavior varies naturally within the population. By contrast, a skewed distribution shows that a behavior is relatively consistent within a population. Following previous research (Kelly et al., 2011) individual infants were deemed to be of interest if their behavior fell +/- 1.5 SDs from the sample mean. We conducted Fisher's Exact Tests for each of the measures to explore the frequency of deviant z-scores in the preliminary HR data relative to the TD sample.

Skewed distributions

Incongruent Responses

The distribution of incongruent responses is heavily skewed (See Figure 5a), with infants consistently displaying a low frequency of incongruent responses. Inspection of the z-scores shows that 2 out of 11 HR infants (18.2 %) produced unusually high frequencies of incongruent responses. By contrast only 8 out of 145 TD infants (5.5%) displayed comparable behavior. A Fisher's Exact Test found that risk status was not significantly associated with a deviant negative z-score (p = .149). Incongruent responses might be a measure of interest for future research, but statistical significance will have to be investigated in a larger sample.

Eye Triggers

The distribution of eye triggers shows a clear skew (See Figure 5b) with infants highly likely to initially fixate the eye area. Inspection of HR infants' z-scores shows that 3 out of 11 HR infants (27.3%) and 13 out of 151 TD infants (8.6%) displayed deviant behavior. A Fisher's Exact Test revealed that risk status was not significantly associated with a deviant negative z-score (*p* = .080). Decreased eye triggers might be a measure of interest for future research, but statistical significance will have to be investigated in a larger sample.

Normal distributions

Eye Region Dwell Time

Dwell time on the eye region is normally distributed (See Figure 5c), indicating that this behavior naturally varies within the population. Within this normal distribution HR z-scores are all located in the left side of the distribution with 4 out of 11 infants (36.4%) showing a negative z-score larger than 1.5. Conversely, only 7 out of 151 TD infants (4.64%) displayed comparable behavior. A Fisher's Exact Test revealed that risk status was significantly associated with a deviant negative z-score (*p* = .003). A relatively low total dwell time on the eye region seems to be characteristic of HR infants of interest. Eye region dwell time when viewing interactive faces should be considered a measure of interest for future research.

Due to our small HR sample, these results are relatively provisional. Although not all Fisher's Exact Tests have reached significance (yet), we observed heightened rates of occurrences in the HR sample comparable to what we should expect based on what previous research tells us about the percentage of HR infants that will eventually receive a diagnosis (~20%, e.g. Ozonoff et al., 2011).

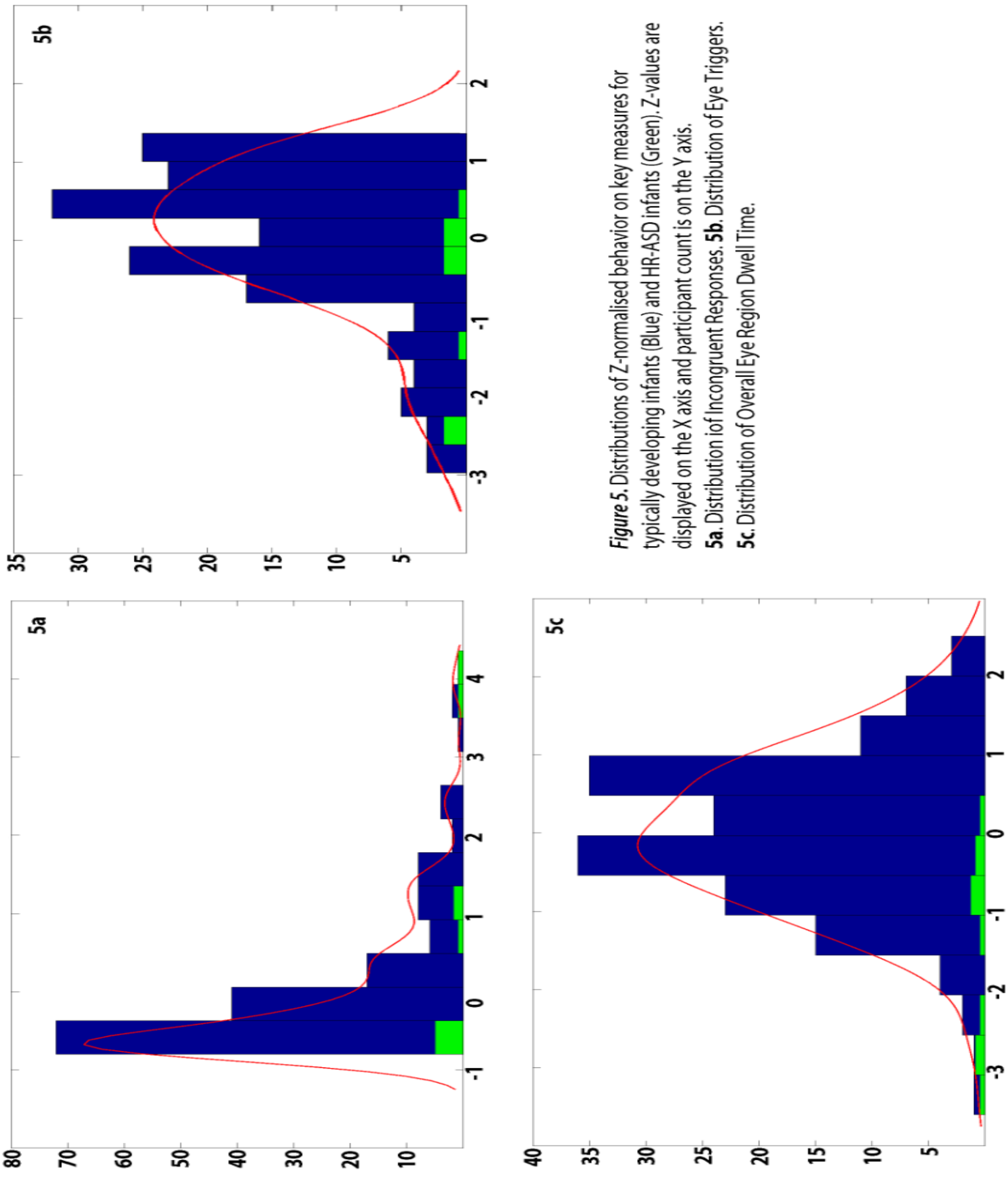


Figure 5. Distributions of Z-normalised behavior on key measures for typically developing infants (Blue) and HR-ASD infants (Green). Z-values are displayed on the X axis and participant count is on the Y axis.
5a. Distribution iof Incongruent Responses. **5b**. Distribution of Eye Triggers.
5c. Distribution of Overall Eye Region Dwell Time.

Figure 5. Distributions of Z-normalized Behavior on Key Measures for Typically Developing Infants (Blue) and High-Risk Infants (Green). Z-values are displayed on the X axis and participant count on the Y axis. **5a.** Distribution of Incongruent Responses. **5b.** Distribution of Eye Triggers. **5c.** Distribution of Overall Dwell Time on Eyes.

Discussion

Previous studies exploring face scanning in infancy have employed *non-interactive* stimuli to answer questions about an inherently *interactive* process. For the current study, we developed a novel eye-tracking method, a GC paradigm, which allowed us to simulate the social context in which face scanning typically occurs in day-to-day life by presenting infants with interactive faces.

Dwell time in a GC paradigm

Our findings demonstrate that regardless of condition infants spent more time fixating the eye area relative to the mouth area, which fits with previous research (e.g. Haith et al., 1977). The mouth area was fixated more in the *SE* condition, which is accounted for by the infants' tendency to look directly at the mouth once a smile was initiated. In accordance with previous studies (e.g. Lewkowicz & Hansen-Tift, 2011) we found that dwell time on eyes declined between 6 and 12-months. Thus, these previously established findings were supported within a GC paradigm. Additionally, there was a relatively large percentage of fixations on other face areas compared to the mouth area (See Supplementary Information for details), which is likely a consequence of the task's interactivity. Scanning a static image of an isolated face in a lab setting could encourage repetition of a triangular pattern of fixations (e.g., right eye – left eye – nose) while disregarding other face areas. In an interactive paradigm, dynamic movement attracts a broader distribution of fixation patterns that is likely to be more representative of natural face-to-face interactions.

Contingency learning

Within social interaction, contingent responses are highly important for the infant's development of social understanding (Markova & Legerstee, 2006) and from as young as 2 months, infants are sensitive to contingencies (e.g. Soussignan et al., 2006). Previous studies have overlooked the contingency of social interaction, whereas our GC stimuli provided this critical element. Consequently, we hypothesized that infants would explicitly learn the contingency of our task and that their subsequent initial fixations would fall in the area of the face that resulted in triggering a socially engaging response. However, we found that infants were more likely to initially fixate the eye area, regardless of condition. More specifically, infants in the *SM* condition did not show evidence of learning that fixating the mouth would produce a socially engaging response. In other words, even when fixating the eye area triggered a socially disengaging response, infants persisted in making eye contact. This replicates previous research demonstrating a strong preference for eye contact (e.g. Di Giorgio et al., 2013). Additionally, this supports the view that infants are deploying a well-rehearsed strategy of engaging in eye contact in order to engage in social interaction and that 20 trials provided insufficient training time to completely deter infants from this behaviour.

Interestingly, infants did show evidence of *implicit* contingency learning, which was inferred by contrasting saccadic response time from trials 1-10 and trials 11-20. Across trials saccadic response times were *decreasing* in the SE condition (i.e. engagement), but *increasing* in the SM condition (i.e., disengagement). Thus, while 20 trials were not enough to demonstrate explicit contingency learning, our GC task was capable of detecting infants' sensitivity to engaging and disengaging actors whilst scanning their faces (See Additional Analyses 2 in Supplementary Information).

Reciprocity

Previous studies exclusively focussed on the infants' eye movements deployed during face scanning. As our interactive paradigm allowed us to simulate a social interaction, we were able to study the additional measure of infants' reciprocity. Infants clearly showed a difference in behavior towards socially engaging and socially disengaging actors. Infants who received a socially engaging response provided a higher frequency of positive responses, suggesting that they enjoyed interacting with the on-screen actor. Conversely, infants who repeatedly triggered a socially disengaging response seemed to withdraw from the task, which is further highlighted by the fact that we had to exclude 16 infants from this condition due to complete disengagement. Infants who did respond to a socially disengaging actor, displayed clear disagreement. Although the overall response rate across conditions appears low (36.17%), it is important to point out that infants were interacting with unfamiliar faces. Relative to previous research on stranger sociability in infancy (e.g. Corter, 1973), our reported response rates are notably high. Our interactive task encouraged infants' active engagement and facilitated responsiveness and sociability (see Ross & Goldman, 1977).

Other applications

Our findings demonstrate that the implementation of GC stimuli allows for a more nuanced investigation of face scanning. We were able to collect measures of contingency and reciprocity, and infants appeared sensitive to social nuances observable in both their eye-tracking (saccadic response times) and behavioral data. As an additional strength, we believe the task could be employed to explore early signs of atypical social development due to a more naturalistic and socially demanding experience. In addition to dwell time, our task can provide measures of contingency learning and reciprocity, skills that are reportedly less developed in children with autism spectrum disorder (ASD) (e.g. Constantino, Przybeck, Darrin & Todd, 2000). To preliminarily investigate this application, we descriptively compared a small sample of infants at high-risk for ASD to z-normalized ranges of typical behavior. Decreased dwell time on eyes seemed to be associated with HR status, which corresponds with earlier findings (Merin et al. 2007). In light of previous studies (e.g. Lewkowicz & Hansen-Tift, 2011), we would expect older infants to redirect their focus to the eye area of a face. Given that the majority of our HR infants were 12 months old, our findings seem to indicate deviant behavior. Additionally, a high frequency of incongruent responses and a lower frequency of eye triggers could be potential measures of interest. A larger HR sample is required to further probe these findings and to assess whether these measures are indeed associated with HR status. Although preliminary, these findings demonstrate that even in a sample with 11 HR infants, a GC paradigm is capable of highlighting HR infants of interest, and suggests potential utility for contributing to early detection.

Conclusion

No studies to date had employed a GC paradigm in face scanning research, whereas the interactive nature of the paradigm lends itself perfectly for research in the area of social interaction. The increased ecological validity of our interactive stimuli permitted us to expand on earlier findings on face scanning by providing measures of contingency learning and reciprocity in addition to a more naturalistic dwell time analysis. Infants clearly showed sensitivity to differences in engagement from actors, which was visible in both saccades and their overt behavioural responses. We preliminarily demonstrated the potential application of a GC paradigm in atypical populations, but further studies are required to corroborate these findings. One limitation to our study was the relatively low power in the Condition x Response interaction for behavioural responses. Accordingly, future studies will require larger sample sizes to address this shortcoming. We believe that when implemented correctly, interactive GC stimuli will allow for more meaningful conclusions in eye-tracking studies in both typical and atypical developmental populations, and will make important contributions to advancements in the field of developmental psychology.

Acknowledgments

This research was supported by a British Academy Grant awarded to Dr. David J. Kelly. We would like to thank the individuals who contributed to the creation of our stimuli, as well as the parents and infants who participated. We would also like to thank Dr. David Meary for supplying our audio-visual calibration targets.

References

Alessandri, S. M., Sullivan, M. W., & Lewis, M. (1990). Violation of expectancy and frustration in early infancy. *Developmental Psychology, 26,* 738-744.

Angulo-Kinzler, R. M., Ulrich, B., & Thelen, E. (2002). Three-Month-Old Infants Can Select Specific Leg Motor Solutions. *Motor Control, 6,* 52-68. doi: 10.1123/mcj.6.1.52

Bushnell, I. W. R., Sai, F., Mullin, J. T. (1989). Neonatal Recognition of the Mother’s Face. *British Journal of Developmental Psychology, 7,* 3-15. doi: 10.1111/j.2044-835X.1989.tb00784.x

Chawarska, K., Macari, S., & Shic, F. (2013). Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with ASD. *Biological Psychiatry, 74,* 195-203. doi: 10.1016/j.biopsych.2012.11.022

Colombo, J., Mitchell, D. W., Coldren, J. T., & Atwater, J. D. (1990). Discrimination learning during the first year: Stimulus and positional cues. *Journal of Experimental Psychology, 16,* 98-109.

Constantino, J. N., Przybeck, T., Friesen, D., & Todd, R. D. (2000). Reciprocal social behavior in children with and without pervasive developmental disorders. *Journal of Developmental and Behavioral Pediatrics, 21*, 2-11. doi: 10.1097/00004703-200002000-00001

Corter, C. M. (1973). A Comparison of the Mother’s and a Stranger’s Control over the Behavior of Infants. *Child Development, 44,* 705-713. doi: 10.2307/1127712

Deligianni, F., Senju, A., Gergely, G., & Csibra, G. (2011). Automated Gaze-Contingent Object Elicit Orientation Following in 8-months-old Infants. *Developmental Psychology, 47*, 1499-1503. doi: 10.1037/a0025659

De Schonen, S., & Bry, I. (1987). Interhemispheric communication of visual learning: A developmental study in 3-6-month old infants. *Neuropsychologia, 25*, 601-612. doi: 10.1016/0028-3932(87)90051-0

Di Giorgio, E., Méary, D., Pascalis, O., & Simion, F. (2013). The Face Perception System Becomes Species-Specific at 3 Months: An Eye-Tracking Study. *International Journal of Behavioral Development, 37,* 95-99. doi: 10.1177/0165025412465362

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye Contact Detection in Humans From Birth. *PNAS, 99,* 9602-9605. doi: 10.1073/pnas.152159999

Fawcett. C., & Liszkowski, U. (2012). Infants Anticipate Others' Social Preferences. *Infant and Child Development, 21,* 239-249. doi: 10.1002/icd.739

Guiraud, J. A., Tomalski, P., Kushnerenko, E., Ribeiro, H., Davies, K., Charman, T., …the BASIS Team. (2012). Atypical Audiovisual Speech Integration in Infants at Risk for Autism. *Plos One, 7,* e36428. doi: 10.1371/journal.pone.0036428

Hains, S. M. J., & Muir, D. W. (1996). Effects of Stimulus Contingency in Infant-Adult Interactions. *Infant Behavior and Development, 19,* 49-61. doi: 10.1016/S0163-6383(96)90043-0

Haith, M. M., Bergman, T., Moore, M. J. (1977). Eye Contact and Face Scanning in Early Infancy. *Science, 198,* 853-855. doi: 10.1126/science.918670

Hauf, P., & Aschersleben, G. (2008). Action-effect anticipation in infant action control. *Psychological Research, 72,* 203-210. doi: 10.1007/s00426-006-0101-3

Hunnius, S., & Geuze, R. H. (2004). Developmental Changes in Visual Scanning of Dynamic Faces and Abstract Stimuli in Infants: A Longitudinal Study. *Infancy, 6,* 231-255. doi: 10.1207/s15327078in0602_5

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. doi:

10.1016/0010-0277(91)90045-6

Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of Visual Orienting in Early Infancy: Contingency Learning, Anticipatory Looking and Disengaging. *Journal of Cognitive Neuroscience, 3,* 335-344. doi: 10.1162/jocn.1991.3.4.335

Kelly, D. J., Duarte, S., Meary, D., Bindemann, M., Jaggie, C., & Pascalis, O. (*under review)*. Infants rapidly detect human faces in complex visual scenes.

Kelly, D. J., Liu, S., Rodger, H., Miellet, S., Ge, L., Caldara, R. (2011). Developmental Cultural Differences in Face Processing. *Developmental Science, 14,* 1176-1184.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual Fixation Patterns During Viewing of Naturalistic Social Situations as Predictors of Social Competence in Individuals With Autism. *Arch Gen Psychiatry, 59,* 809-816. doi: 10.1001/archpsyc.59.9.809

Langton, S. R. H., Law, A. S., Burton, A. M., & Schweinberger, S. R. (2008). Attention capture by faces. *Cognition, 107,* 330-342. doi: 10.1016/j.cognition.2007.07.012

Lewkowicz, D. J., & Hansen-Tift, A. M. (2011). Infants deploy selective attention to the mouth of a talking face when learning speech. *PNAS, 109,* 1431-1436. doi: 10.1073/pnas.1114783109

Markova, G., & Legerstee, M. (2006). Contingency, Imitation, and Affect Sharing: Foundations of Infants' Social Awareness. *Developmental Psychology, 42,* 132-141. doi: 10.1037/0012-1649.42.1.132

Maurer, D., & Salapatek, P. (1976). Developmental Changes in the Scanning of Faces by Young Infants. *Child Development, 47*, 523-527. doi: 10.1111/1467-8624.ep12190048

Merin, N., Young, G. S., Ozonoff, S., & Rogers, S. J. (2007). Visual Fixation Patterns during

Reciprocal Social Interaction Distinguish a Subgroup of 6-Month-Old Infants At-Risk for Autism from Comparison Infants. *Journal of Autism and Developmental Disorders, 37*, 108-121. doi: 10.1007/s10803-006-0342-4

Miyazaki, M., Takahashi, H., Rolf, M., Okada, H., & Omori, T. (2014). The image-scratch paradigm: a new paradigm for evaluating infants' motivated gaze control. *Scientific Reports, 4,* 5498. doi: 10.1038/srep05498

Mondloch, C. J., Lewis, T. L., Budreau, D. R., Maurer, D., Dannemiller, J. L., Stephens, B. R., & Kleiner-Gathercoal, K. A. (1999). Face Perception During Early Infancy. *Psychological science, 10,* 419-422. doi: 10.1111/1467-9280.00179

Oakes, L. M., & Ellis, A. E. (2013). An eye-tracking investigation of developmental changes in infants' exploration of upright and inverted human faces. *Infancy, 18,* 134-148. doi: 10.1111/j.1532-7078.2011.00107.x

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., …Stone, W.L. (2011). Recurrence Risk for Autism Spectrum Disorders: A Baby Siblings Research Consortium Study. *Pediatrics, 128,* 488-495. doi: 10.1542/peds.2010-2825

Pascalis, O., De Schonen, S., Morton, J., Deruelle, C., & Fabre-Grenet, M. (1995). Mother's Face Recognition by Neonates: A Replication and an Extension. *Infant Behavior and Development, 18,* 79-85. doi: 10.1016/0163-6383(95)90009-8

Rochat, P., Querido, J. G., Striano, T. (1999). Emerging sensitivity to the timing and structure of protoconversation in early infancy. *Developmental Psychology, 35,* 950-957. doi: 10.1037/0012-1649.35.4.950

Ross H. S., & Goldman, B. D. (1977). Infants' Sociability toward Strangers. *Child Development, 48*, 638-642. doi: 10.2307/1128666

Rovee, C. K., & Rovee, D. T. (1969). Conjugate reinforcement of infant exploratory behaviour. *Journal of Experimental Child Psychology, 8*, 33-39. doi: 10.1016/0022-0965(69)90025-3.

Shic, F., Macari, S., & Chawarska, K. (2014). Speech Disturbs Face Scanning in 6-Month-Old Infants Who Develop Autism Spectrum Disorder. *Biological Psychiatry, 75*, 231-237. doi: 10.1016/j.biopsych.2013.07.009

Slater, A., Von der Schulenburg, C., Brown, E., Badenoch, M., Butterworth, G., & Parsons, S. (1998). Newborn Infants Prefer Attractive Faces. *Infant Behavior and Development, 21,* 345-354. doi: 10.1016/S0163-6383(98)90011-X

Soussignan, R., Nadel, J., Canet, P., & Gerardin, P. (2006). Sensitivity to Social Contingency and Positive Emotion in 2-Month-Olds. *Infancy, 10,* 123-144. doi: 10.1207/s15327078in1002_2

Tenenbaum, E. J., Shah, R. J., Sobel, D. M., Malle, B. F., Morgan, J. L. (2012). Increased Focus on the Mouth Among Infants in the First Year of Life: A Longitudinal Eye-Tracking Study. *Infancy, 18,* 534-553. doi: 10.1111/j.1532-7078.2012.00135.x

Valenza, E., Simion, F., Macchi Cassia, V., Umilta, C. (1996). Face Preference at Birth. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 892-903. doi: 10.1037/0096-1523.22.4.892

Wagner, J. B., Luyster, R. J., Tager-Flusberg, H., & Nelson, C. A. (2016). Greater Pupil Size in Response to Emotional Faces as an Early Marker of Social-Communicative Difficulties in Infants at High Risk for Autism. *Infancy, 21*, 560-581. doi: 10.1111/infa.12128

Wang, Q., Bolhuis, J., Rothkopf, C. A., Kolling, T., Knopf, M., Triesch, J. (2012). Infants in

Control: Rapid Anticipation of Action Outcomes in a Gaze-Contingent Paradigm. *PLoS ONE, 7,* e30884. doi: 10.1371/journal.pone.0030884.

Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G. R., & Vogely, K. (2010). It’s in your Eyes – Using Gaze-contingent Stimuli to Create Truly Interactive Paradigms for Social Cognitive and Affective Neuroscience. *Social Cognitive and Affective Neuroscience, 5*, 98-107. doi: 10.1093/scan/nsq024

Young, G. S., Merin, N., Rogers, S. J., & Ozonoff S. (2009). Gaze Behavior and Affect at 6-Months: Predicting Clinical Outcomes and Language Development in Typically Developing Infants and Infants At-Risk for Autism. *Developmental Science, 12,* 798-814. doi: 20.2222/j.2467-7687.2009.00833.x

Legends for tables and figures

Table 1. Participant Characteristics per Age group and Condition

Figure 1. A Stimulus with the Discrete Eye and Mouth Regions Visible. The individual whose face appears here gave signed consent for her likeness to be published in this article.

Figure 2. Examples of Socially-Engaging and Socially-Disengaging Animations with Accompanying Infants' Behavioral Responses. The authors received signed consent for the woman's and children's likenesses to be published in this article.

Figure 3. Summed fixations post-animation onset for all infants in the *SE* condition on the left and *SM* condition on the right. The tendency for infants to fixate the mouth (smile) in the *SE* condition but not in the *SM* condition is clearly visible. The individual whose face appears here gave signed consent for her likeness to be published in this article.

Figure 4. Average Saccadic Latencies for Both Conditions on Trials 1-10 and Trials 11-20. Between conditions there was no difference in saccadic latencies for trials 1-10, but there was a significant difference for trials 11-20, implicating implicit contingency learning. Within conditions there was a significant difference in saccadic latencies between trials 1-10 and trials 11-20 for only the *SE* condition.

Figure 5. Distributions of Z-normalized Behavior on Key Measures for Typically Developing Infants (Blue) and High-Risk Infants (Green). Z-values are displayed on the X axis and participant count on the Y axis. **5a.** Distribution of Incongruent Responses. **5b.** Distribution of Eye Triggers. **5c.** Distribution of Overall Dwell Time on Eyes.