



Kent Academic Repository

Kelly, David J., Duarte, Sofia, Meary, David, Bindemann, Markus and Pascalis, Olivier (2019) *Infants rapidly detect human faces in complex naturalistic visual scenes*. *Developmental Science*, 22 (6). ISSN 1363-755X.

Downloaded from

<https://kar.kent.ac.uk/73096/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://doi.org/10.1111/desc.12829>

This document version

Author's Accepted Manuscript

DOI for this version

Licence for this version

UNSPECIFIED

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal*, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).



DR. DAVID KELLY (Orcid ID : 0000-0002-8033-7556)

Article type : Paper

Infants rapidly detect human faces in complex naturalistic visual scenes

Running title

Infant face detection

Authors

David J. Kelly¹, Sofia Duarte¹, David Meary², Markus Bindemann¹, & Olivier Pascalis²

Author affiliations

¹School of Psychology, Keynes College, University of Kent, Canterbury, Kent, CT2 7NP

2. Universite Grenoble Alpes, Grenoble, France

Corresponding Author

David J. Kelly, School of Psychology, Keynes College, University of Kent, Canterbury, Kent, CT2

7NP

Email: d.j.kelly@kent.ac.uk; Telephone:

Conflict of Interest

There are no conflicts of interest to declare.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/desc.12829

This article is protected by copyright. All rights reserved.

Acknowledgements

The research team would like to thank Jonathan E. Prunty, Jolie R. Keemink and Kelsey Jackson for their contributions to data collection. We thank Daniel LaCombe for providing an earlier version of Matlab code used for the RQA analysis. We also thank all the caregivers and infants who kindly contributed to this study.

Research highlights

- First ever comprehensive demonstration of infant's rapid orienting to human faces in complex naturalistic visual scenes
- In contrast to previous findings using artificial 'circular-array presentations', evidence for a face bias was found in infants aged 6 month and younger
- A range of novel oculomotor and scene analysis comparisons including the first implementation of recurrence quantification analysis on an infant eye-tracking dataset

Abstract

Infants respond preferentially to faces and face-like stimuli from birth, but past research has typically presented faces in isolation or amongst an artificial array of competing objects. In the current study infants aged 3- to 12-months viewed a series of complex visual scenes; half of the scenes contained a person, the other half did not. Infants rapidly detected and oriented to faces in scenes even when they were not visually salient. Although a clear developmental improvement was observed in face detection and interest, all infants displayed sensitivity to the presence of a person in a scene, by displaying eye movements that differed quantifiably across a range of measures when viewing scenes that either did or did not contain a person. We argue that infant's face detection capabilities are ostensibly 'better' with

naturalistic stimuli and artificial array presentations used in previous studies have underestimated performance.

Keywords: face detection, infancy, eye movements, visual search

Human faces represent a unique class of stimulus (e.g., Kanwisher, 2000) that attract greater visual interest relative to other stimulus groups (Langton, Law, Burton, & Schweinberger, 2008), including non-human primate faces (Pascalis, de Haan, & Nelson, 2002; Pascalis & Kelly, 2009). Newborn infants display a visual interest in faces and preferentially orient towards face-like stimuli when presented alongside a competing stimulus (e.g., Johnson, Dziurawiec, Ellis, & Morton, 1991; Mondloch et al., 1999; Valenza, Simion, Cassia, & Umiltà, 1996). Furthermore, a third trimester foetus will preferentially head turn towards a simple upright 3-point face configuration light display, but not an inverted counterpart (Reid et al., 2017). Although much is known about the developmental trajectory of face recognition, categorisation and preference (See Pascalis et al., 2011 for a review), there is a relative paucity of research investigating face detection; here defined as the ability to rapidly align the fovea with a face located in the visual periphery. This action is ubiquitous in everyday visual behaviour with face detection likely to be subcortical in origin and involving the superior colliculus, pulvinar and amygdala (de Gelder, Frissen, Barton, & Hadjikhani, 2003; Johnson, 2005; Johnson, Senju, & Tomalski, 2015). Attending to faces allows us to extract social information (Todorov, Said, Engell, & Oosterhof, 2008) and assess threat from our visual environment (LoBue & DeLoache, 2010; Ohman, 2005) while simultaneously disregarding other visually salient information that might otherwise attract attention (Itti & Koch, 2000). A visual

interest in faces and eye contact is also thought to be critical for successful social development (Farroni, Csibra, Simion, & Johnson, 2002).

Adults can detect faces in just 100 msec (Crouzet, Kirchner, & Thorpe, 2010) and evidence supports the view that faces represent a distinct class of visual stimuli (Mackay, Cerf, & Koch, 2012) processed by a dedicated neural architecture (Haxby, Hoffman, & Gobbini, 2000). In infants, it is believed that a low-spatial frequency face detection system, termed *Conspec*, is subcortical, and exists from birth through to adulthood (Johnson, et al., 2015; Morton & Johnson, 1991). Face detection theories postulate and promote the importance of direct gaze in the subcortical route (Stein, Peelen, & Sterzer, 2011) and suggest that eye contact facilitates face detection (Senju & Johnson, 2009). Convergent data from behavioural and neuroanatomical studies advocate the presence of a subcortical route in newborn infants, which serves to detect faces and to orient the newborn towards them (Johnson, 2005). Yet, despite the strong evidence for an interest in faces from birth, a recent review of the infant face detection literature concluded that a general face bias is weak in early infancy, but becomes more robust from 6 months of age onwards (Leppanen, 2016). However, previous studies have predominantly used simplified circular-array visual presentations, which have yielded mixed findings. Face detection is reported for colour images (Elsabbagh et al., 2013; Gliga, Elsabbagh, Andravizou, & Johnson, 2009; Gluckman & Johnson, 2013; Kwon, Setoodehnia, Baek, Luck, & Oakes, 2016), but not for grayscale images (Di Giorgio, Turati, Altoe, & Simion, 2012). Additionally, a lone visual preference task assessing spontaneous orienting to faces versus toys found no advantage for faces (DeNicola, Holt, Lambert, & Cashon, 2013). A more complex 5 x 5 visual grid display was utilised to show efficient face *finding* (i.e. looking to the face in the display; Jakobsen, Umstead, & Simpson, 2016), but *detection* was not assessed.

Although these findings provide insight into infant face detection, faces are rarely viewed in isolation in non-laboratory conditions, they are never seen embedded within an array (circular or grid) nor are they seen at a constant distance. Instead, recent studies using head-mounted camera recording methods have confirmed that faces are typically viewed alongside substantial levels of competing visual information and at myriad distances (Jayaraman, Fausey, & Smith, 2017; Sugden, Mohamed-Ali, & Moulson, 2014). Consequently, the utility of artificial presentations and the generalisability of findings from existing work is unclear. Studies presenting infants with more realistic visual scenes (Amso, Haas, & Markant, 2014; Frank, Amso, & Johnson, 2014) have reported increasing attention to faces in the first year of life and beyond, but face detection has not been assessed using such stimuli. While eye movements in visual search using naturalistic scenes have long been studied in adults (e.g., Rayner, 2009), comparable work in infants is lacking. Wass and Smith (2014) have provided a detailed description of eye movement behaviour during a range of dynamic scene viewing tasks, but other studies with infant populations have been largely restricted to visual array (Kwon, et al., 2016) or other non-naturalistic visual presentations (Hessels, Hooge, & Kemner, 2016).

It is important to note that although one may have the subjective impression that we see *all information* in our visual field, detailed object analyses is unlikely to be conducted over a wide area in the visual periphery (Henderson & Hollingsworth, 1998) and we can only overtly attend to (i.e. foveate) a single location at a time. Indeed, change blindness (e.g., Simons & Levin, 1997) provides powerful evidence to support the view that we do not actively process all details present in our visual periphery. However, in adults at least, peripheral vision does play a critical role in determining the location of upcoming fixations. Adults seemingly extract global 'gist' information about scenes (e.g., 'sea', 'mountain' etc.) with presentations as brief as 26 ms (Rousselet, Joubert, & Fabre-Thorpe, 2005) and are able to rapidly assess the presence of a

human regardless of scene variance (Rousselet, Mace, & Fabre-Thorpe, 2003). It is further believed that extracting gist and scene layout can help to guide attention to likely target locations based on prior knowledge and task instruction (e.g., Navalpakkam & Itti, 2005). Assessing what infants find to be visually salient (i.e. the perceptual qualities that make certain items 'stand out' relative to other items and consequently attract attention) is more challenging though as researchers must ask questions *without* language and prior knowledge is necessarily limited in infant participants. Accordingly, the extent to which these findings from the adult literature can be applied to infants is currently unclear.

Determining what attracts visual attention during scene viewing in adults is well-studied with seminal studies of eye movements showing that certain discrete locations, such as human form, are fixated more than 'other' information (Buswell, 1935; Yarbus, 1961; 1967). In order to account for eye movements patterns, early models of visual salience (e.g., Itti & Koch, 2000) promoted the role of bottom-up information in determining shifts of covert attention while more recent models incorporate top-down information in the form of prior knowledge and task instruction (e.g., Navalpakkam & Itti, 2005). Indeed, it has been demonstrated, for example, that adult's fixations are not accounted for by saliency when viewing scenes containing social information (Birmingham, Bischof, & Kingstone, 2009). However, with restricted top-down input in the first year of life, as noted above, it is uncertain whether such models extend to infant visual attention. Indeed, assessment of salience in predicting infant eye movements is conflicting with an increase in attention to salience across the first year of life reported by some authors (Althaus & Mareschal, 2012; Amso, et al., 2014) and a decrease in attention to salience reported by others (Kwon, et al., 2016).

In response to calls for a need to shift towards more naturalistic stimuli in infant research (Frank, Vul, & Johnson, 2009) and to disentangle the influences of bottom-up and top-down influences on attention and learning (Althaus & Mareschal, 2012), the current study aims

to explore spontaneous face detection in naturalistic scenes, while simultaneously considering the role of visual salience in guiding eye movements. We report data collected from a total of 241 infants comprising four separate age groups and eye movements recorded at 500Hz with low data loss (~8% samples during on-screen looking). Thus, we explored the ability of infants aged 3-, 6-, 9- and 12-months to detect visually salient and non-salient faces embedded within complex, naturalistic scenes. Infants viewed four scenes that contained a person (2 salient, 2 non-salient) and four scenes that did not contain a person. To assess sensitivity to faces within scenes, we distinguish between two behaviours: *detecting* and *finding*. A face was considered *detected* if the first saccadic movement post-stimulus onset was directed to the face.

Additionally, we also included instances when infant's saccades undershot or overshot the stimulus (maximum 1° visual angle) and a corrective saccade was made (within 200 msec) to foveate the face on the next fixation. A face was considered *found* if it was fixated at any point during the 5-second trial. If a face was not fixated, it was coded as *missed* (See Figure 1). For the sake of brevity, *face finding* and subsequent interest in faces was extremely high for all age groups (suggesting a strong bias for faces at all ages) and results are reported in the Supplementary Online Materials (SOM). This manuscript will therefore focus on our primary interests, which are *face detection* and eye movements during scene processing with a particular emphasis on *Person Present* and *Person Absent* trials. Following findings in adults (Birmingham, et al., 2009), it could be predicted that infants will detect faces regardless of visual salience. However, if infants are less influenced by top-down control as a consequence of no task instruction and their limited visual experience with the world, we can expect eye movements to be better predicted by salience, especially in the youngest age groups. In addition to this explicit contrast, we will also report a range of exploratory analyses designed to capture any fine-grained differences in eye movement behaviour that are elicited by the presence or absence of a person in a visual scene.

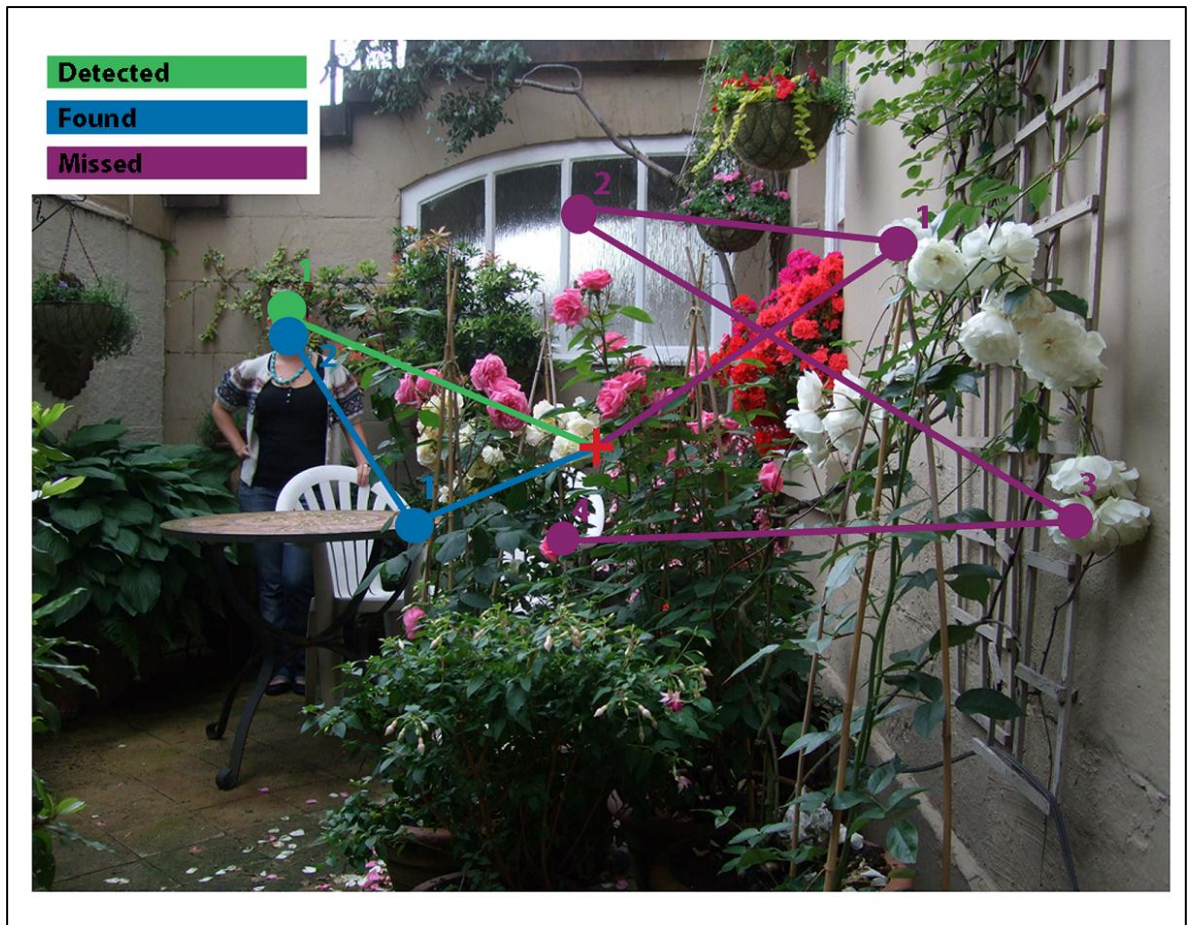


Figure 1. Schematised scanpaths and fixations for examples of *detected*, *found* and *missed* faces.

Method

Participants

At the University of Kent participants were contacted via the Kent Child Development Unit database following initial recruitment at local mother and baby groups. At the Universite Grenoble Alpes, infants were recruited directly from the local maternity hospital. Infants were deemed eligible to participate if they were within a +/- 14 day age range of a target age at the time of testing. Infants with any known visual impairments were ineligible for the study. The

Accepted Article

participant's caregivers were provided with an information sheet prior to testing and additionally given the opportunity to verbally ask questions before signing a consent form. Participants and caregivers were compensated with age-appropriate gifts. The study received ethical approval from both institutions.

A total of 241 infants were included in the analyses with a further 22 infants excluded for failing to produce useable data (3 months, $n = 1$; 6 months, $n = 6$; 9 months, $n = 8$; 12 months, $n = 7$). Infants were omitted from the final sample for failing to complete all 8 trials ($n = 14$) or providing unanalysable data ($n = 8$) as a consequence of extreme movement. This final sample comprised infants from 4 distinct age groups; 3 months ($n = 23$), 6 months ($n = 65$), 9 months ($n = 82$) and 12 months ($n = 71$; See Table S4). All Infants were randomly allocated to Image Condition A or Image Condition B (See Figure S1).

Stimuli

Each stimulus was analysed for its visual content using the Visual Saliency Toolbox (Walther & Koch, 2006). The toolbox assesses the low-level visual properties (e.g., brightness, contrast etc.) of images and produces a map that highlights the most visually salient stimulus properties (see Figure 2 and SOM for further details). Preliminary visual saliency analyses were initially conducted on a large pool of images ($n = 120$) with sixteen images containing the desired properties selected for the final stimulus set. Following the approach of Amso and colleagues (Amso, et al., 2014) we identified images in which the person either was or was not considered to be a visually salient aspect of the scene. All images were photographs of complex indoor and outdoor scenes that had been used in a previous study (Bindemann, Scheepers, Ferguson, & Burton, 2010). There were two versions of each scene; one with a *Person Present* and one with a *Person Absent*. In half of the *Person Present* images ($n = 4$), part of the pixel

space occupied by the person was determined to be visually salient by the toolbox, whereas in the remaining images ($n = 4$) the pixel space occupied by the person was not visually salient.

The *Person Present* images were divided evenly between stimulus groups, such that each infant was presented with two salient and two non-salient *Person Present* images. The set of sixteen images were separated into two stimulus sets (Image Conditions A and B), with each containing eight images: four *Person Present* and four *Person Absent*. Replication of scenes within each stimulus set was avoided. As infants participated in just one condition, they saw every individual scene once only.

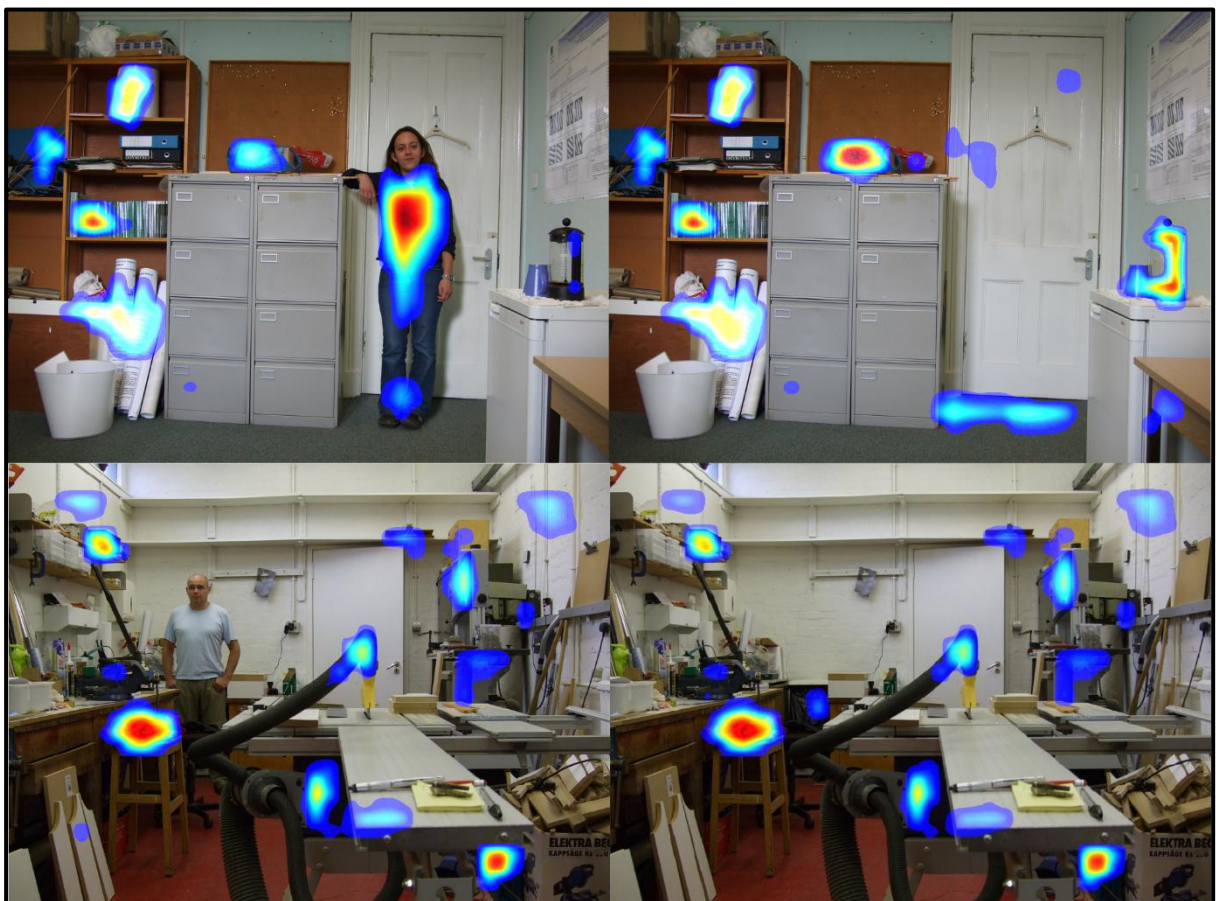


Figure 2. Examples of 'Person present' and 'Person absent' versions of two different scenes.

Heat Map data representing Visual Saliency as computed by the Visual Saliency Toolbox

(Walther & Koch, 2006) are overlaid. Zero saliency values have been made transparent and peak saliency is shown in red, fading to blue. The person in the top left panel was categorised as 'Salient' and the person in the bottom left panel was categorised as 'non-salient'.

Materials

Eye movements were recorded with an EYELINK 1000+ (SR Research, Ontario) at a sampling rate of 500 Hz operated in Head Reference 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.

The stimuli were presented using Experiment Builder (SR Research, Ontario, CA) and the raw eye movement data were extracted using Data Viewer (SR Research). Fixations and saccades were subsequently parsed in Matlab (The Mathworks, MA, USA) using custom written code (See SOM for full details). All subsequent data processing was completed using further custom written Matlab code.

Procedure

The caregivers of the participants were greeted and taken to a waiting room. After signing the consent form, caregivers and infant participants were escorted to the research

laboratory. Infants were placed in an age-appropriate padded seat in front of a computer monitor positioned at a distance of 60 centimetres. Testing was conducted in low light conditions. In order to operate the Eyelink in Remote Mode, a small target sticker was placed centrally on the infant's forehead. The target serves as an external reference point to the tracked eye. The infant's right eye was tracked throughout testing. The infant's view to their surroundings, caregiver/s and experimenters was obstructed by an occluding screen in order to minimize distractions. A 5-point calibration procedure using custom-made attention-grabbing audio-visual targets was conducted initially and repeated as many times as required. No infant failed to calibrate. Following successful calibration (calibration-validation error $< 1^\circ$), the task was immediately initiated. The eight test images were presented sequentially for 5 seconds each in a fully randomised order. An attention grabber appeared in the centre of the screen between each stimulus presentation that centred the infant's gaze for the beginning of each trial. The trial was initiated only when the infant was fixating the target ($< 1^\circ$), so accordingly the infant was fixating the screen centre at the start of each trial.

Results

Preliminary analyses revealed no significant differences across testing sites, between Image Conditions A and B or participant gender in terms of saccadic velocities, number of fixations or fixation duration, so data were collapsed for further analyses. Additionally, following previous work (e.g., Hessels, et al., 2016; Hooge & Erkelens, 1996) first fixations were excluded from fixation duration analysis. We report analyses related to Face Detection and Interest, Oculomotor Control and *Person Present vs Person Absent* scene comparison. The goals of the reported analyses are to explore how infants respond when viewing complex visual

scenes that do or do not contain a social stimulus (i.e. a person) and to use a variety of analysis methods to capture potential eye movement differences during scene viewing.

Face Detection

In order to assess face detection rates, paired-samples *t*-tests contrasted the proportion of trials in which faces were *detected* (from *Person Present* trials) against the proportion of trials in which the first saccade was directed towards salient areas of equated pixel space (from *Person Absent* trials; see Figure S2). In order to achieve this, we constructed salience AOI maps that contained the most visually salient areas as computed by the Saliency Toolbox. All ages groups were significantly more likely to direct a first saccade towards a face than towards a salient AOI ($t_3(22) = 4.467$; $t_6(64) = 16.153$; $t_9(82) = 19.233$; $t_{12}(71) = 27.080$; all $ps < .001$; See Tables S5 - S6 and SOM for further details and additional comparisons).

Detecting Faces: Saliency Effects

In order to assess whether infant's detection of faces was influenced by their visual salience, a series of 4 (Age) X 2 (Saliency: Salient or Non-Salient) mixed ANOVAs were conducted to explore the effect of saliency on *Face Detection*, *Face Detection Saccade Latency* and *Face Dwell Time*, which was defined as the proportion of total time spent looking at a face after it had been initially fixated (see Table 1 for a summary of means). For Face Detection and Face Dwell Time, all infants were included in the analyses, but a total of 13 infants did not detect any faces (3 months, $n = 9$; 6 months, $n = 2$; 9 months, $n = 2$) and were consequently omitted from the Face Detection Saccade Latency analysis.

Detection: Salient faces were more likely to be detected relative to non-salient faces ($F(1, 237) = 9.171$, $p = .003$, $\eta_p^2 = .037$) with improvements in *Age* also found ($F(3, 237) =$

27.506, $p < .001$, $\eta_p^2 = .258$), although the interaction failed to reach significance ($F(3, 237) = 1.197$, $p = .312$, $\eta_p^2 = .015$).

Face Detection Saccade Latency: The ANOVA yielded a main effect of *Age* ($F(3, 224) = 11.090$, $p = .001$, $\eta_p^2 = .129$) only with *post-hoc* comparisons showing significant differences between 3- and 12-month-olds only ($p < .03$). The main effect of *Salience* and *Age x Salience* interaction did not reach significance.

Face Dwell Time: The ANOVA yielded a main effect of *Salience* ($F(1, 237) = 81.929$, $p < .001$, $\eta_p^2 = .263$), a main effect of *Age* ($F(3, 237) = 18.894$, $p < .001$, $\eta_p^2 = .198$) and a *Salience x Age* interaction ($F(3, 237) = 5.571$, $p < .001$, $\eta_p^2 = .068$). To explore the age-related differences, a *post-hoc* one-way ANOVA yielded significant *Age* differences for *Salient Faces* ($F(3, 237) = 5.939$, $p < .001$, $\eta_p^2 = .070$) with further comparisons showing significant differences for 3 month-olds versus all other age groups only ($p_s < .001$). A further one-way ANOVA yielded significant *Age* differences for *Non-Salient Faces* ($F(3, 237) = 21.966$, $p < .001$, $\eta_p^2 = .223$) with significant differences for all age group comparisons.

Table 1. Mean Number of Salient and Non-Salient faces detected per age group (95% CIs in parenthesis)

| Age | Salient Face Trials | | | Non-Salient Face Trials | | |
|------------------|--------------------------|---------------------------------------|--------------------------------|--------------------------|---------------------------------------|--------------------------------|
| | Faces Detected (max = 2) | Face Detection Saccade Latency (Secs) | Face Dwell Time (Proportional) | Faces Detected (max = 2) | Face Detection Saccade Latency (Secs) | Face Dwell Time (Proportional) |
| 3 months | .48 [0.25; 0.69] | .83 [0.68; 0.97] | .50 [0.44; 0.56] | .43 [0.22; 0.65] | .87 [0.52; 1.22] | .23 [0.18; 0.28] |
| 6 months | 1.29 [1.06; 1.40] | .58 [0.44; 0.73] | .65 [0.58; 0.72] | .80 [0.71; 1.01] | .61 [0.52; 0.71] | .36 [0.28; 0.44] |
| 9 months | 1.31 [1.14; 1.46] | .42 [0.36; 0.48] | .66 [0.60; 0.72] | 1.15 [1.02; 1.29] | .46 [0.42; 0.51] | .52 [0.46; 0.58] |
| 12 months | 1.54 [1.39; 1.68] | .40 [0.37; 0.44] | .73 [0.66; 0.79] | 1.33 [1.19; 1.48] | .44 [0.40; 0.48] | .61 [0.56; 0.66] |

Individual Image Analysis

In order to determine whether age-related differences in detection were related to any potential differences in oculomotor control, all available face detection peak saccadic velocity data was entered into a one-way ANOVA (i.e. data from all trials when the face was successfully detected), which revealed no age-related differences ($F(3, 555) = .587, p = .624, \eta p^2 = .003$). However, the saccadic amplitudes required to *Detect* a face (i.e. movement from

Accepted Article

the screen centre to the face) differed across images (see Figure 3 Top Panel and Table S7). Accordingly, Saccadic Amplitudes and Peak Saccadic Velocity were highly correlated ($r = .968$, $n = 559$, $p < .001$) with a clearly observable linear relationship present (See Figure 3 bottom left panel) that is consistent with the well-reported relationship between amplitude and velocity termed the *main sequence* (e.g., Bahill, Clark & Stark, 1975; Hainline, Turkel, Abramov, Lemerise, & Harris, 1984). Any observable differences in face detection behaviour are unlikely to stem from oculomotor control deficits between age groups and instead will reflect a failure to detect a person in a scene as a consequence of improvements in visual acuity, which are known to occur across the age ranges tested (Courage & Adams, 1990; Dobson & Teller, 1978; Sokol, 1978). To further explore *Image* related differences, a one-way ANOVA was conducted on saccadic amplitudes, revealing highly significant differences ($F(7, 551) = 824.649$, $p < .001$, $\eta_p^2 = .913$). *Post hoc* comparisons found significant differences between all *Image* comparisons except between *Images* 1 – 3, which had the most closely matched screen centre to face amplitudes. Images varied in terms of difficulty as highlighted by differences in detection rates between images $\chi^2(7, N = 950) = 161.246$, $p < .001$ and differences in saccadic onset time ($F(3, 555) = 8.945$, $p < .001$, $\eta_p^2 = .046$) with both measures highly negatively correlated ($r = -.959$, $n = 8$, $p < .001$; see Figure 3 bottom right panel).

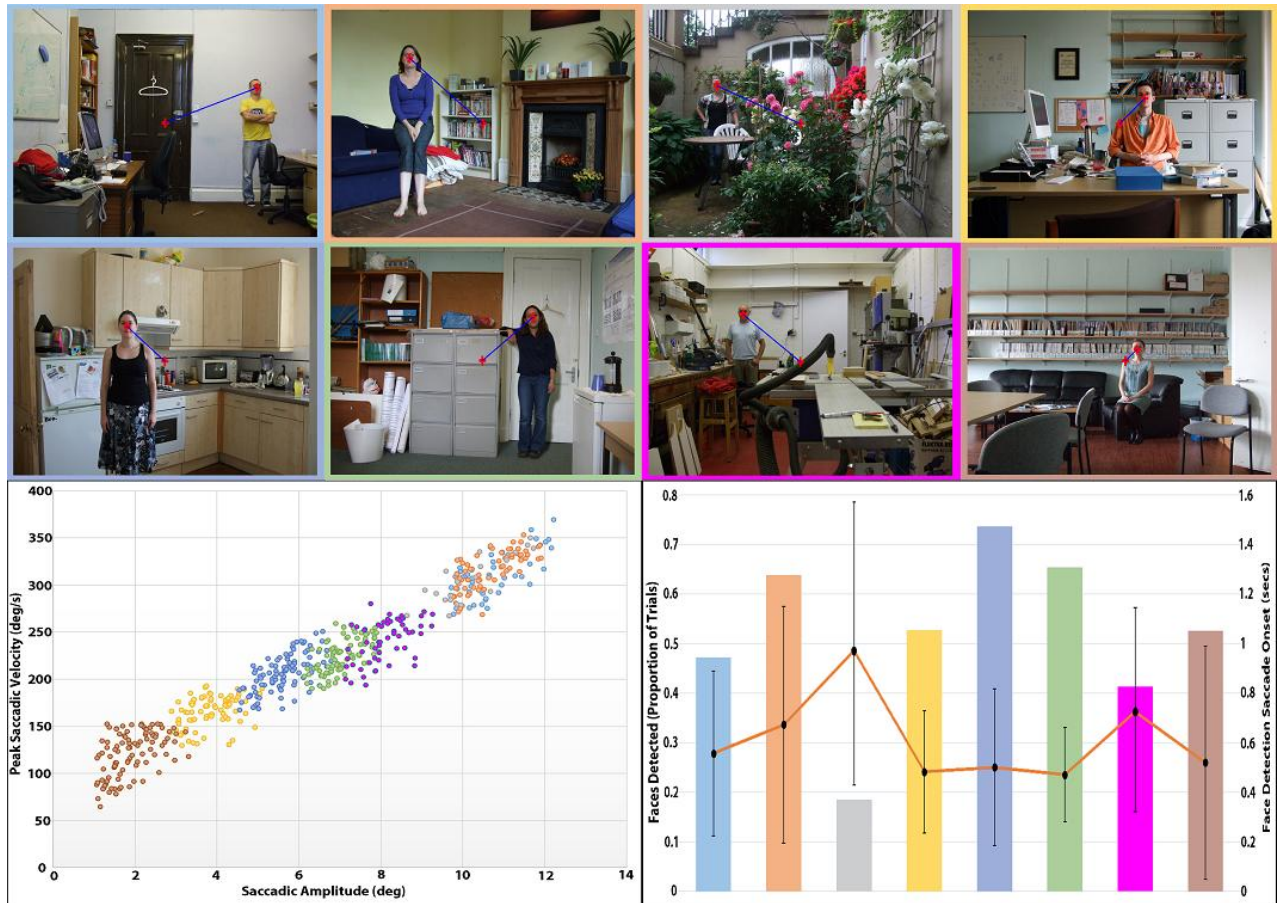


Figure 3. Top: Colour coded *Person Present* images (each image has an individual colour that corresponds to data in the bottom panels) with schematised saccadic movement required for detection; Bottom Left: Saccadic Amplitude and Peak Velocity plot for all detection saccades; Bottom Right: Proportion of trials in which detection occurred with corresponding mean saccade latencies.

Person Present vs Person Absent Trials Contrast: Scene Processing

To explore additional measures of scene processing, we next present analyses that investigate how the presence of a person impacts on eye movement behaviour. To explore behavioural differences in eye movements for trials containing and not containing people, a

series of A 4 (*Age*) X 2 (*Person*: Present or Absent) mixed ANOVAs were conducted on the number of fixations and fixation durations.

A clear pattern emerged from the analyses (See Figure S3 and SOM for more details). Relative to *Person Absent* scenes, *Person Present* scenes received more overall looking (i.e. summed fixation durations) ($F(1, 237) = 13.903, p < .001, \eta_p^2 = .052$) and they received fewer fixations ($F(1, 237) = 7.316, p = .007, \eta_p^2 = .030$), but they were of a longer duration ($F(1, 237) = 30.199, p < .001, \eta_p^2 = .113$). Finally, a 4 (*Age*) X 2 (*Fixation Location*: Face or Other (i.e. all non-face locations) repeated measures ANOVA conducted on *Fixation Location* found that all age groups deployed fixations of significantly longer durations ($F(1, 237) = 139.980, p < .001, \eta_p^2 = .371$) when fixating the *Face* ($M = 701$ msec) relative to *Other* locations ($M = 447$ msec). Additionally, the distribution of fixation durations was also markedly different (See Figure 4 and Table S8).

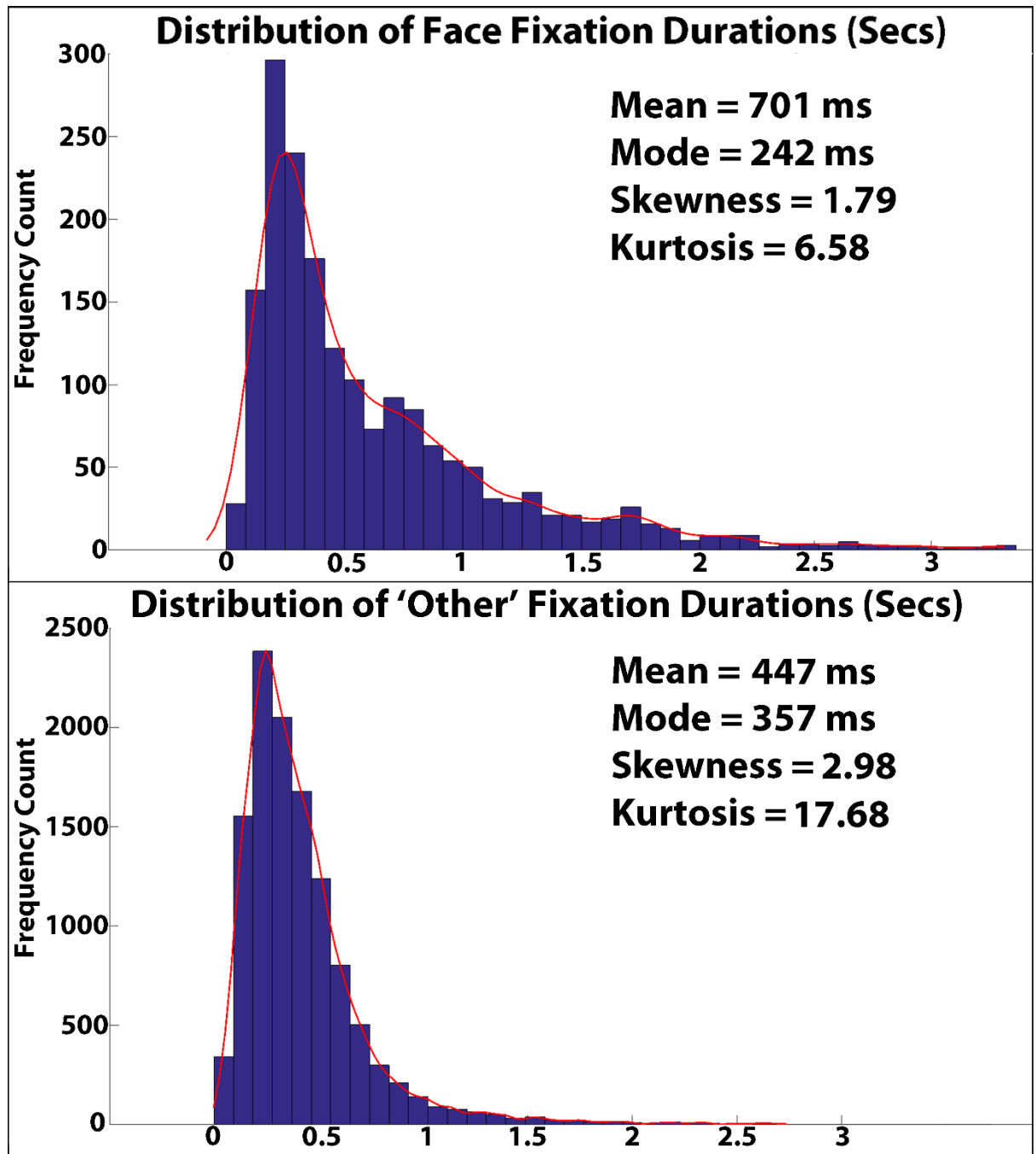


Figure 4. Distribution of fixation durations landing on the face and 'other' locations.

Recurrence Quantification Analysis

To further characterise fixation sequences, recurrence quantification analysis (RQA) was conducted. RQA has previously been used to describe complex dynamic systems such as climatological data (Marwan, Wessel, Meyerfeldt, Schirdewan, & Kurths, 2002), but more recently it has successfully been applied as a tool for analysing spatial and temporal aspects of eye movement behaviour (Anderson, Bischof, Laidlaw, Risko, & Kingstone, 2013). In RQA, the closeness of two fixations f_i and f_j is measured using a fixed radius p . Fixations are considered recurrent if they are spatially close as defined by their Euclidean distance d , so that fixations are recurrent if $d(f_i, f_j) \leq p$. In principal, one can select any radius size, but recurrence rates will necessarily escalate as a function of increasing size. For the current experiment, a radius of 2° of visual angle (60 pixels) was selected to approximate the fovea. RQA analysis was implemented in Matlab using code provided by Daniel LaCombe.

Recurrence describes how frequently observers refixate discrete locations.

Determinism represents the proportion of gaze patterns that are repeated.

Laminarity describes the likelihood for discrete regions of a scene to be repeatedly fixated.

Centre of Recurrence Mass (CORM) indicates whether refixations are temporally close or distant.

Smaller values represent refixations occurring closely in time.

See Anderson et al. (2013) for further computational details, figures, algorithms and discussion on the merits of *fixation-distance methods* vs. *fixed-grid methods* (e.g., ScanMatch; Cristino, Mathot, Theeuwes, & Gilchrist, 2010).

The four RQA measures were analysed separately with 4 (*Age* (months): 3, 6, 9, 12) X 2 (*Person*: present or absent) mixed ANOVAs. Following Anderson et al. (2013), trials that contained no recurrence ($n = 206$) were removed prior to analysis. This equates to 10.8% of total trials ($n = 1898$).

The primary findings were main effects of *Person* found for *Recurrence* ($F(1, 1684) = 47.265, p < .001, \eta_p^2 = .027$), *Determinism* ($F(1, 1684) = 4.699, p = .030, \eta_p^2 = .003$), *Laminarity* ($F(1, 1684) = 16.793, p < .001, \eta_p^2 = .010$) and *CORM* ($F(1, 1684) = 10.521, p = .001, \eta_p^2 = .006$) with higher rates found in *Person Present* trials relative to *Person Absent* trials for each (See SOM for main effects of *Age* and interactions).

The results from the RQA analysis require prudent interpretation. Recurrence should be observed in scene viewing (Anderson et al. 2013), so accordingly the recurrence rates observed in *Face absent* scenes are not unexpected. However, the elevated rates of recurrence reported for *Person Present* represent a quantitative shift in fixation patterns that are supported by the more classical analyses (reported above) that contrasted *Person Present* and *Person Absent* trials. In brief, stimuli that contain faces capture attention and consequently fewer fixations occur, but they are often longer and almost exclusively directed to the face (See Figure S3a). In addition, discrete areas (i.e the face) are scanned in finer detail as evidenced by the laminarity results. Stimuli that do not contain faces are subjected to visual exploration by the infant, which produces a greater number of shorter duration fixations dispersed over a broader spatial area (See Figure S3b).

Discussion

Infants of all age groups consistently detected faces embedded within complex visual scenes regardless of their visual salience, despite the fact that the presence and spatial location of a person followed no systematic pattern, they subtended a small visual angle, sclera was inconspicuous, they were presented alongside substantial competing visual information and the actual face of a person in a scene did not constitute the most visually salient aspect of a scene (See Figure 2). Overall, we believe that this is the clearest, most comprehensive, and

explicit demonstration of face detection ever reported in an infant population. The sample sizes tested in the current study greatly exceed ‘typical’ sample sizes for infant research and allow us to draw conclusions with confidence.

To some extent, our findings contrast with conclusions drawn from previous research that suggested a bias for faces is weak in the first 6 months of life (Leppanen, 2016). Although we report a developmental improvement in detection and interest in faces, it is important to acknowledge that visual acuity develops substantially between 3- and 6-months of age (e.g., Courage & Adams, 1990) and this is likely to have contributed to the difference in performance. Additionally, the *face finding* rate in 3-month-old infants was 2.6 out of 4 (See SOM), which is suggestive of a bias for faces at this age. By 6-months of age infants are displaying a clear bias for faces. We propose that our stimuli can account for the lack of early bias reported previously. Specifically, we assert that the naturalistic scenes we used provided a closer approximation of the infant’s real-world visual experiences (see Jayaraman, et al., 2017 for examples of infant head-camera mounted stills) than visual grid/array displays used previously (cf. Di Giorgio, et al., 2012; Gliga, et al., 2009; Gluckman & Johnson, 2013; Jakobsen, et al., 2016) and helped to facilitate detection. One might intuitively assume that complex visual scenes would make face detection more difficult relative to visual arrays/grids. However, naturalistic scenes provide context for the objects that they contain (e.g., Biederman, Mezzanotte, & Rabinowitz, 1982) and in our specific case, at least some of the person’s body was present in each image, which presumably also provided some contextual information for infants and may have helped to guide their target selection. It is critical to note that grid and array presentations do not provide such context and accordingly this might account for the ‘weak face bias’ that has been reported previously.

The individual image analysis provided clear evidence that saccadic velocity is comparable with adult’s saccadic parameters and found no obvious age-related improvements

from 3- to 12-months of age. By contrast, face detection saccadic latencies did differ significantly between groups with delayed onsets observed in younger infants relative to older infants. We propose that younger infants simply take longer to process the image / extract gist information and consequently longer saccadic latencies are observed. It is, however, worth noting that although there was substantial overlap in the ranges of saccadic latencies produced by all age groups, the 3-month old age participants displayed fewer instances of face detection and consequently provided fewer saccadic latency data points to assess. Accordingly, a larger sample of 3-month-old infants will be required in future studies to fully explore this issue. Additionally, while our data demonstrate rapid detection of faces, future research should look to identify the precise mechanisms that underpin this ability. In some instances, face detection saccadic latencies were as short as 150 msec, which are nearly as fast as latencies reported in adults (Crouzet, et al., 2010) and are faster than would be permitted by cortical systems (Johnson, et al., 2015; Schmolesky et al., 1998). Our findings are also potentially at odds with the view that eye-gaze might be necessary to activate the subcortical route (Senju & Johnson, 2009; Stein, et al., 2011). Although each person appearing in our stimuli had their eyes open, relative to previous work (Farroni, et al., 2002), the sclera was not a prominent feature. It is difficult to discount the role of eye contact in supporting the face detection we report, but our findings suggest that an empirical study designed to address this question explicitly would be highly informative.

The inclusion of salience as a factor in our analyses enabled us to contrast our findings with past studies in adults (e.g., Birmingham, Bischof, & Kingstone, 2009) and to assess the relative contributions of bottom-up and top-down attentional control. While this topic has been studied and debated for some decades in the adult literature, a comparable literature with infant participants does not exist. Although we report evidence of detection and interest in faces for salient and non-salient faces, we did find significantly better performance with

salient faces suggesting that visual salience might play a role in guiding eye movements. However, while we followed the approach of Amso and colleagues (Amso, et al., 2014) to categorise people as ‘salient’ or ‘non-salient’ it is important to note that this is a relatively arbitrary dichotomy that does not truly reflect the visual world. Further studies will help to determine the applicability of existing saliency models to infant vision and to evaluate the relative contributions of top-down and bottom-up control in driving visual attention across ontogeny. Additionally, the spatial and temporal dynamics of infant eye movements are currently under-studied and warrant further investigation. We have demonstrated here that RQA analyses can be performed effectively with an infant dataset and is capable of capturing fine-grained differences in performance that are not detectable using classic ‘AOI’ analysis only. However, infant’s eye movement behaviour during scene processing represents a largely untapped research area (but see Helo, Rama, Pannasch, & Meary, 2016) that warrants further empirical investigation.

In addition to face detection, we have shown that the mere presence of a person in a visual scene has a measurable quantitative impact on eye movement behaviour. Adult’s fixation durations during scene viewing have a skewed distribution with a mode of 230 ms and a mean of 330 ms (Henderson & Hollingworth, 1999). It is notable that infant’s fixations directed to ‘other’ locations closely mirrors this pattern of behaviour with a mode of 357 ms and mean of 447 ms. Based on findings from EEG studies (e.g., de Haan, Pascalis, & Johnson, 2002), it is reasonable to speculate that these differences simply reflect infant’s slower processing. Interestingly, fixation durations to faces are also skewed but with a mode of 242 ms and a mean of 701 ms. We suggest that this disparity is explained by the fact that infants make some very long initial fixations to faces (> 1000 ms), but also produce a large quantity of short duration fixations deployed during fine-grained scanning as highlighted by the significant *laminarity* results reported above.

In conclusion, we have provided clear evidence that infants aged 3- to 12-months rapidly attend to discrete faces embedded within complex visual scenes. We have characterised scanning behaviour to further emphasise the importance of faces to infants and how their presence impacts eye movement behaviour across several dependent measures. We suggest that future studies should endeavour to depart from using artificial visual arrays and instead try to use more naturalistic stimuli that are both more familiar to infants and are more likely to provide a genuine reflection of their capabilities.

References

- Althaus, N., & Mareschal, D. (2012). Using saliency maps to separate competing processes in infant visual cognition. *Child Dev*, *83*(4), 1122-1128. doi: 10.1111/j.1467-8624.2012.01766.x
- Amso, D., Haas, S., & Markant, J. (2014). An eye tracking investigation of developmental change in bottom-up attention orienting to faces in cluttered natural scenes. *PLoS One*, *9*(1), e85701. doi: 10.1371/journal.pone.0085701
- PONE-D-13-36576 [pii]
- Anderson, N. C., Bischof, W. F., Laidlaw, K. E., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. *Behav Res Methods*, *45*(3), 842-856. doi: 10.3758/s13428-012-0299-5
- Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: detecting and judging objects undergoing relational violations. *Cogn Psychol*, *14*(2), 143-177. doi: 0010-0285(82)90007-X [pii]
- Bindemann, M., Scheepers, C., Ferguson, H. J., & Burton, A. M. (2010). Face, body, and center of gravity mediate person detection in natural scenes. *J Exp Psychol Hum Percept Perform*, *36*(6), 1477-1485. doi: 10.1037/a0019057
- 2010-16533-001 [pii]
- Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision Res*, *49*(24), 2992-3000. doi: 10.1016/j.visres.2009.09.014
- S0042-6989(09)00432-5 [pii]
- Courage, M. L., & Adams, R. J. (1990). Visual acuity assessment from birth to three years using the acuity card procedure: cross-sectional and longitudinal samples. *Optom Vis Sci*, *67*(9), 713-718.
- Cristino, F., Mathot, S., Theeuwes, J., & Gilchrist, I. D. (2010). ScanMatch: a novel method for comparing fixation sequences. *Behav Res Methods*, *42*(3), 692-700. doi: 10.3758/BRM.42.3.692
- 42/3/692 [pii]
- Crouzet, S. M., Kirchner, H., & Thorpe, S. J. (2010). Fast saccades toward faces: face detection in just 100 ms. *J Vis*, *10*(4), 16 11-17. doi: 10.1167/10.4.16

10/4/16 [pii]

de Gelder, B., Frissen, I., Barton, J., & Hadjikhani, N. (2003). A modulatory role for facial expressions in prosopagnosia. *Proc Natl Acad Sci U S A*, *100*(22), 13105-13110. doi: 10.1073/pnas.1735530100

1735530100 [pii]

de Haan, M., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *J Cogn Neurosci*, *14*(2), 199-209. doi: 10.1162/089892902317236849

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*(2), 143-147. doi: 10.1177/0165025412474751

Di Giorgio, E., Turati, C., Altoe, G., & Simion, F. (2012). Face detection in complex visual displays: an eye-tracking study with 3- and 6-month-old infants and adults. *J Exp Child Psychol*, *113*(1), 66-77. doi: 10.1016/j.jecp.2012.04.012

S0022-0965(12)00081-1 [pii]

Dobson, V., & Teller, D. Y. (1978). Visual acuity in human infants: a review and comparison of behavioral and electrophysiological studies. *Vision Res*, *18*(11), 1469-1483. doi: 0042-6989(78)90001-9 [pii]

Elsabbagh, M., Gliga, T., Pickles, A., Hudry, K., Charman, T., & Johnson, M. H. (2013). The development of face orienting mechanisms in infants at-risk for autism. *Behav Brain Res*, *251*, 147-154. doi: 10.1016/j.bbr.2012.07.030

S0166-4328(12)00493-7 [pii]

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proc Natl Acad Sci U S A*, *99*(14), 9602-9605. doi: 10.1073/pnas.152159999

152159999 [pii]

Frank, M. C., Amso, D., & Johnson, S. P. (2014). Visual search and attention to faces during early infancy. *J Exp Child Psychol*, *118*, 13-26. doi: 10.1016/j.jecp.2013.08.012

S0022-0965(13)00178-1 [pii]

Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, *110*(2), 160-170. doi: 10.1016/j.cognition.2008.11.010

S0010-0277(08)00260-6 [pii]

Gliga, T., Elsabbagh, M., Andravizou, A., & Johnson, M. (2009). Faces Attract Infants' Attention in Complex Displays. *Infancy*, *14*(5), 550-562. doi: 10.1080/15250000903144199

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

10.3389/Fpsyg.2013.00527

Hainline, L., Turkel, J., Abramov, I., Lemerise, E., & Harris, C. M. (1984). Characteristics of saccades in human infants. *Vision Res*, *24*(12), 1771-1780. doi: 0042-6989(84)90008-7 [pii]

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn Sci*, *4*(6), 223-233. doi: S1364-6613(00)01482-0 [pii]

Helo, A., Rama, P., Pannasch, S., & Meary, D. (2016). Eye movement patterns and visual attention during scene viewing in 3- to 12-month-olds. *Vis Neurosci*, *33*, E014. doi: 10.1017/S0952523816000110

S0952523816000110 [pii]

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annu Rev Psychol*, *50*, 243-271. doi: 10.1146/annurev.psych.50.1.243

Hessels, R. S., Hooge, I. T., & Kemner, C. (2016). An in-depth look at saccadic search in infancy. *J Vis*, *16*(8), 10. doi: 10.1167/16.8.10

2528617 [pii]

- Hooge, I. T., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Percept Psychophys*, *58*(7), 969-976.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10-12), 1489-1506. doi: Doi 10.1016/S0042-6989(99)00163-7
- Jakobsen, K. V., Umstead, L., & Simpson, E. A. (2016). Efficient human face detection in infancy. *Dev Psychobiol*, *58*(1), 129-136. doi: 10.1002/dev.21338
- Jayaraman, S., Fausey, C. M., & Smith, L. B. (2017). Why are faces denser in the visual experiences of younger than older infants? *Dev Psychol*, *53*(1), 38-49. doi: 10.1037/dev0000230
- 2016-61509-003 [pii]
- Johnson, M. H. (2005). Subcortical face processing. *Nat Rev Neurosci*, *6*(10), 766-774. doi: nrn1766 [pii]
- 10.1038/nrn1766
- Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1-2), 1-19. doi: 0010-0277(91)90045-6 [pii]
- Johnson, M. H., Senju, A., & Tomalski, P. (2015). The two-process theory of face processing: modifications based on two decades of data from infants and adults. *Neurosci Biobehav Rev*, *50*, 169-179. doi: 10.1016/j.neubiorev.2014.10.009
- S0149-7634(14)00259-0 [pii]
- Kanwisher, N. (2000). Domain specificity in face perception. *Nat Neurosci*, *3*(8), 759-763. doi: 10.1038/77664
- Kwon, M. K., Setoodehnia, M., Baek, J., Luck, S. J., & Oakes, L. M. (2016). The development of visual search in infancy: Attention to faces versus salience. *Dev Psychol*, *52*(4), 537-555. doi: 10.1037/dev0000080
- 2016-07264-001 [pii]
- Langton, S. R., Law, A. S., Burton, A. M., & Schweinberger, S. R. (2008). Attention capture by faces. *Cognition*, *107*(1), 330-342. doi: S0010-0277(07)00204-1 [pii]
- 10.1016/j.cognition.2007.07.012
- Leppanen, J. M. (2016). Using Eye Tracking to Understand Infants' Attentional Bias for Faces. *Child Dev Perspect*, *10*(3), 161-165. doi: 10.1111/cdep.12180
- CDEP12180 [pii]
- LoBue, V., & DeLoache, J. S. (2010). Superior detection of threat-relevant stimuli in infancy. *Developmental Science*, *13*(1), 221-228. doi: 10.1111/j.1467-7687.2009.00872.x
- Mackay, M., Cerf, M., & Koch, C. (2012). Evidence for two distinct mechanisms directing gaze in natural scenes. *J Vis*, *12*(4), 9. doi: 10.1167/12.4.9
- 9 [pii]
- 12.4.9 [pii]
- Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., & Kurths, J. (2002). Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *Phys Rev E Stat Nonlin Soft Matter Phys*, *66*(2 Pt 2), 026702. doi: 10.1103/PhysRevE.66.026702
- 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*(5), 419-422. doi: Doi 10.1111/1467-9280.00179
- Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: a two-process theory of infant face recognition. *Psychol Rev*, *98*(2), 164-181.
- Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Res*, *45*(2), 205-231. doi: S0042-6989(04)00392-X [pii]
- 10.1016/j.visres.2004.07.042

- Ohman, A. (2005). The role of the amygdala in human fear: Automatic detection of threat. *Psychoneuroendocrinology*, 30(10), 953-958. doi: 10.1016/j.psyneuen.2005.03.019
- Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science*, 296(5571), 1321-1323. doi: 10.1126/science.1070223
296/5571/1321 [pii]
- Pascalis, O., de Martin de Vivies, X., Anzures, G., Quinn, P. C., Slater, A. M., Tanaka, J. W., & Lee, K. (2011). Development of face processing. *Wiley Interdiscip Rev Cogn Sci*, 2(6), 666-675. doi: 10.1002/wcs.146
- Pascalis, O., & Kelly, D. J. (2009). The Origins of Face Processing in Humans: Phylogeny and Ontogeny. *Perspect Psychol Sci*, 4(2), 200-209. doi: 10.1111/j.1745-6924.2009.01119.x
4/2/200 [pii]
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Q J Exp Psychol (Hove)*, 62(8), 1457-1506. doi: 10.1080/17470210902816461
911217562 [pii]
- Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017). The Human Fetus Preferentially Engages with Face-like Visual Stimuli. *Curr Biol*, 27(13), 2052. doi: S0960-9822(17)30773-X [pii]
10.1016/j.cub.2017.06.036
- Rousselet, G. A., Joubert, O. R., & Fabre-Thorpe, M. (2005). How long to get to the "gist" of real-world natural scenes? *Visual Cognition*, 12(6), 852-877. doi: 10.1080/13506280444000553
- Rousselet, G. A., Mace, M. J., & Fabre-Thorpe, M. (2003). Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *J Vis*, 3(6), 440-455. doi: 10.1167/3.6.5
3/6/5 [pii]
- Schmolsky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., & Leventhal, A. G. (1998). Signal timing across the macaque visual system. *J Neurophysiol*, 79(6), 3272-3278. doi: 10.1152/jn.1998.79.6.3272
- Senju, A., & Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends Cogn Sci*, 13(3), 127-134. doi: 10.1016/j.tics.2008.11.009
S1364-6613(09)00019-9 [pii]
- Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends Cogn Sci*, 1(7), 261-267. doi: 10.1016/S1364-6613(97)01080-2
S1364-6613(97)01080-2 [pii]
- Sokol, S. (1978). Measurement of infant visual acuity from pattern reversal evoked potentials. *Vision Res*, 18(1), 33-39. doi: 0042-6989(78)90074-3 [pii]
- Stein, T., Peelen, M. V., & Sterzer, P. (2011). Adults' Awareness of Faces Follows Newborns' Looking Preferences. *PLoS One*, 6(12). doi: ARTN e29361
10.1371/journal.pone.0029361
- Sugden, N. A., Mohamed-Ali, M. I., & Moulson, M. C. (2014). I spy with my little eye: typical, daily exposure to faces documented from a first-person infant perspective. *Dev Psychobiol*, 56(2), 249-261. doi: 10.1002/dev.21183
- Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, 12(12), 455-460. doi: 10.1016/j.tics.2008.10.001
- Valenza, E., Simion, F., Cassia, V. M., & Umiltà, C. (1996). Face preference at birth. *J Exp Psychol Hum Percept Perform*, 22(4), 892-903.
- Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Netw*, 19(9), 1395-1407. doi: S0893-6080(06)00215-2 [pii]
10.1016/j.neunet.2006.10.001

Wass, S. V., & Smith, T. J. (2014). Individual Differences in Infant Oculomotor Behavior During the Viewing of Complex Naturalistic Scenes. *Infancy*, 19(4), 352-384. doi: 10.1111/infa.12049

Yarbus, A. L. (1961). Eye movements during the examination of complicated objects. *Biofizika*, 6(2), 52-56.