doi:10.1093/scan/nst012 SCAN (2014) 9, 538–543

What you see is what you get: contextual modulation of face scanning in typical and atypical development

Mayada Elsabbagh,[1] Rachael Bedford,[2] Atsushi Senju,[1] Tony Charman,[2] Andrew Pickles,[3] Mark H. Johnson,[2] and The BASIS Team[2]

[1]Department of Psychiatry, Faculty of Medicine, McGill University, Montreal, Canada, [2]Centre for Brain and Cognitive Development, Birkbeck, University of London, London, UK, and [3]King's College London, Institute of Psychiatry, London, UK

Infants' visual scanning of social scenes is influenced by both exogenously and endogenously driven shifts of attention. We manipulate these factors by contrasting individual infants' distribution of visual attention to the eyes relative to the mouth when viewing complex dynamic scenes with multiple communicative signals (e.g. peek-a-boo), relative to the same infant viewing simpler scenes where only single features move (moving eyes, mouth and hands). We explore the relationship between context-dependent scanning patterns and later social and communication outcomes in two groups of infants, with and without familial risk for autism. Our findings suggest that in complex scenes requiring more endogenous control of attention, increased scanning of the mouth region relative to the eyes at 7 months is associated with superior expressive language (EL) at 36 months. This relationship holds even after controlling for outcome group. In contrast, in simple scenes where only the mouth is moving, those infants, irrespective of their group membership, who direct their attention to the repetitive moving feature, i.e. the mouth, have poorer EL at 36 months. Taken together, our findings suggest that scanning of complex social scenes does not begin as strikingly different in those infants later diagnosed with autism.

Keywords: infant; autism; language development; eye tracking

Human infant's preferential attention to socially relevant information, such as faces from a very early age has been the focus of several theoretical models of typical and atypical development (Johnson *et al.*, 2005). Manipulation of stimuli presented in various studies has allowed further specification of the key characteristics of faces preferentially attracting attention (Farroni *et al.*, 2006). These preferences are robust in the face of manipulation of low-level perceptual, e.g. contrast polarity, illumination and motion features of the scenes. Infant eye-tracking studies demonstrated that infants under 2 months tend to fixate mainly around the edge of the face (Maurer and Salapatek, 1976; Haith *et al.*, 1977). From 2 months, and similarly among adults (Yarbus, 1967), infants begin to fixate on the internal features of the face, such as eyes and mouth. Infants as young as 6 weeks show a strong preference for the internal features of the face when they are watching their mother's face that demonstrate highly communicative expression, such as maintained eye contact, smiling, speaking in infant-directed speech and nodding (Hunnius and Geuze, 2004). It has also been suggested that infants' preferential tracking of the eyes relative to the mouth is reflective of different language acquisition milestones, with interest to the mouth in dynamic scenes being strongest between 4 and 8 months (Lewkowicz and Hansen-Tift, 2012).

Our primary aim in this study is to investigate the origins, and the later developmental consequences of variability in face scanning both in typical and atypical development. Our approach builds on a number of eye-tracking studies of scanning of social stimuli in individuals with autism spectrum disorders (henceforth, autism or ASD). Atypical use of eye contact to regulate social interaction is among the defining clinical features of autism. An influential claim in this area has been that differences in scanning of social scenes reflect, or may indeed lead to, the range of social and communication impairments characterizing the condition. For example, some eye-tracking studies have revealed that individuals with autism fixate others' eyes less than typically developing individuals (Klin *et al.*, 2002; Pelphrey *et al.*, 2002). However, other studies failed to replicate this pattern (van der Geest *et al.*, 2002; Dapretto *et al.*, 2006) or reported mixed results (Neumann *et al.*, 2006; Speer *et al.*, 2007). These findings have generated competing hypotheses with some researchers suggesting that less looking towards the eyes relative to the mouth predicts more severe autism symptoms, whereas others have proposed that increased looking towards the mouth is a compensatory mechanism reflected in a reduction in communication symptoms (Senju and Johnson, 2009).

In tracing the developmental origins of these putative face scanning differences in the autism phenotype, we motivate our study on the basis of well-established developmental models that have demonstrated that the infant in the first year is an active and efficient forager of environmental input in general (Robertson *et al.*, 2004), with increased attention to potential social communicative situations in particular (Csibra and Gergely, 2009; Gliga and Csibra, 2009). Specifically, we consider individual differences in the ability to modulate attention in response to a complex and varying environment as reflecting variation in 'endogenous control'. The latter is defined as variation in infants' ability to exert control over their own looking behaviour, irrespective of conflicting demands for attention from the environment (Johnson, 1990). Endogenous control is often contrasted with exogenous control, where attention is driven reflexively by external events. It is largely accepted that the two orienting mechanisms rely on overlapping neural architecture, but experimental studies can manipulate the extent to which endogenous mechanisms are engaged relative exogenous ones (Johnson, 1990). For example, the degree to which endogenous mechanisms of attention are engaged in extracting socially relevant from complex stimuli has been previously studied in typical individuals (Langdell, 1978; Deaner and Platt, 2003; Nummenmaa and Calder, 2009). Such manipulation often relies on manipulating the social context, its complexity and/or other task demands. In individuals with autism, such factors have a profound impact on performance.

Received 13 March 2012; Accepted 27 January 2013
Advance Access publication 5 February 2013

We are very grateful for the enormous contributions BASIS families have made towards this study. The research is supported by The UK Medical Research Council (G0701484) and the BASIS funding consortium led by Autistica (www.basisnetwork.org) to M.H. Johnson. Further support for some of the authors is from Autism Speaks and COST action BM1004. The BASIS team in alphabetical order: Simon Baron-Cohen, Patrick Bolton, Susie Chandler, Janice Fernandes, Holly Garwood, Teodora Gliga, Kristelle Hudry, Greg Pasco, Leslie Tucker and Agnes Volein.

Correspondence should be addressed to Mayada Elsabbagh, PhD, Assistant Professor, Department of Psychiatry, Faculty of Medicine, McGill University, 1033 Pine Avenue West Montreal, Quebec H3A 1A1, CANADA. E-mail: mayada.elsabbagh@mcgill.ca

Reduced fixations on the eyes is most commonly reported with complex and cognitively demanding face stimuli, e.g. by obscuring faces with 'Bubbles' masks (Neumann *et al.*, 2006; Spezio *et al.*, 2007) or by using dynamic stimuli (Klin *et al.*, 2002; Speer *et al.*, 2007; Riby and Hancock, 2009). Several behavioural studies also report that individuals with ASD rely less on the upper part of the face when they process faces (Langdell, 1978; Joseph and Tanaka, 2003). As such, context-sensitive modulation of looking behaviour is likely to reflect endogenous influences on visual selection.

Despite mixed findings, variable visual scanning profiles in autism, which are most likely related to atypical endogenous control, appear to map onto some aspects of the condition. Children with ASD whose socio-emotional behaviours are relatively less impaired than their non-verbal communication look more at the eyes, whereas those with the opposite profile look more at the mouth (Falck-Ytter *et al.*, 2010). Other studies have suggested that these difficulties in face scanning explain a wider range of impairments in processing of other face information and more generally, socially relevant information. For example, the duration of spontaneous fixation on the eyes correlates with the level of activation in fusiform gyrus (Dalton *et al.*, 2005) and specific instruction to fixate the eyes results in the typical level of activation in fusiform gyrus (Hadjikhani *et al.*, 2004; Hadjikhani *et al.*, 2007) in individuals with ASD. Interestingly, similar results were observed in a group of siblings of individuals with autism who do not themselves have a diagnosis (Dalton *et al.*, 2007). It has been suggested that studying individual variability among infants at familial risk for autism may provide a powerful approach by extending the range of variability in outcomes observed in typical development (Elsabbagh and Johnson, 2010).

It is also important to note that these putative differences in endogenous control in autism may also depend on the developmental stage of the individual. For example, Chawarska and Shic (2009) showed that reduced fixation on the eyes in ASD, previously suggested as a characteristic of autism, may not be present at a younger age. In their longitudinal study, 2-year-old children with autism showed similar fixation to the eyes as typically developing children, even though they showed less fixation on the mouth. At 4 years of age, children with autism spent less time looking at the inner parts of the face including eyes, mouth and nose than typically developing children. However, the difference in the amount of fixation to the eyes between groups did not reach significance.

Taken together, previous studies suggest potential developmental differences in endogenous control of attention in autism which is evident in that (i) individuals with autism differ from control groups in context-dependent modulation of fixation patterns, (ii) variation in fixation patterns appears to map onto different symptom profiles seen in the condition and (iii) such differences emerge over time through dynamic developmental pathways. Nevertheless, direct evidence for developmental accounts based on studies with much younger infants, has been lacking given that autism is rarely diagnosed in the first 2 years of life. Yet, the presence of atypical eye contact in early development could potentially hamper a wide range of social learning, as eye contact is known to play a critical role in communicative learning (Csibra and Gergely, 2009; Senju and Johnson, 2009). For example, in typical development, preferential orienting to eye contact is present even in newborns (Farroni *et al.*, 2002). Atypical eye contact processing may also contribute to a range of social and communicative symptoms commonly observed in young children with autism (Loveland and Landry, 1986; Charman *et al.*, 2003). Yet, the apparent lack of differences in looking towards the eye in toddlers with autism seem to be inconsistent with this account (Chawarska and Shic, 2009).

Our previous studies designed to examine infants' exogenously driven orienting to faces suggest that infants later diagnosed with autism do not vary in their reflexive orienting to faces embedded within a simple static array of distractors (Elsabbagh *et al.*, 2013). To date, however, only one prospective study has tested the longitudinal correspondence between early face scanning and later autism-related outcomes using tasks that engage more endogenous relative to exogenous control. Typically developing 6-month-old infants looked equally to the eyes and mouth when interacting with an adult, but the infants increased fixations of the eyes relative to mouth in the 'Still-face' period, during which the adult suddenly froze, became expressionless and stopped interacting with the infants (Merin *et al.*, 2007). In the same study, infants at familial risk of autism did not differ as clearly in their scanning but a small subgroup of infants at risk looked more to the mouth relative to the eyes. A follow-up study with a larger group (including the infants in the previous study) found that more mouth relative to eyes fixations did not relate specifically to a later diagnosis of autism/ASD, but did relate to individual differences in expressive language (EL) as assessed at 24 months (Young *et al.*, 2009).

Taken together, these studies provide key lessons. First, while specific regions of the face, namely the eyes, may attract infants' attention in complex scenes, endogenous control mechanisms enable the infant to flexibly reorient attention to other regions. Second, variability in dynamic scanning observed early in life may reflect, or even lead to, specific developmental outcomes. Third, rather than having a specific imbalance in attention to the eyes as compared with the mouth, individuals with autism may exhibit differences in the balance of exogenous and endogenous factors modulating their attention to socially relevant information embedded within complex dynamic stimuli.

In this study, we attempted to integrate these different considerations in a unified design and within a large group of typically developing infants, and in infants at increased familial risk for developing autism by virtue of having an older sibling with a diagnosis of the condition. The latter group is one where we expect significantly variable profiles in the development of social and communication skills, which at the extreme may manifest in an autism diagnosis (Elsabbagh and Johnson, 2010). In previous studies, we used orienting paradigms to examine exogenous *vs* endogenous orienting using well-controlled simple scenes, but ones that are impoverished relative to the infants' natural social environment (e.g. Elsabbagh *et al.*, 2009, 2011; Holmboe *et al.*, 2010). In this study, we tested contextual modulation of the relationship between early eye-tracking measures and later developmental outcomes. More specifically, we contrasted the infants' scanning patterns of a familiar and socially rich scene of *peek-a-boo* that engages more endogenous mechanisms, with simpler scenes in which different features on the face moved independently (eyes-, mouth- or hand-moving) and are therefore less likely to engage endogenous but more likely to engage exogenous attentional mechanisms. As such, we examined infants' performance in *peek-a-boo* scenes that combine multiple communicative features relative to their performance when each feature was manipulated independently. The inclusion of infants at risk in our study may shed light on any atypical mechanisms associated with atypical developmental outcomes and/or emerging characteristics of autism.

METHODS

Participants

One hundred and four infants from the British Autism Study of Infant Siblings took part in this study (54 at risk, 21 male and 50 low risk, 21 male). Along with several other measures, the infants participated in the eye-tracking task when they were 6- to 10-months old and again when they were 12–15 months. Subsequently, 52 (from 54) of those at risk for ASD were seen for assessment around the second birthday and 53 around their third birthday by an independent team. During the 36-month visit, a battery of clinical research measures was administered (see Supplementary materials for details). Consensus ICD-10 criteria were used to ascertain diagnosis in a subgroup of infants at risk using all available information from all visits by experienced researchers (TC, KH, SC and GP). Supplementary materials present detailed participant characteristics, such as ascertainment of risk status, background measures at each visit and outcome characterization including clinical classification. The at-risk group were classified as having ASD (Sib-ASD), other developmental concerns (Sib-Other) or typically developing (Sib-Typical).

Eye-tracking study at 6–10 months and 12–15 months

During their first and second visits, infants were administered a battery of eye-tracking tasks containing non-identical stimuli across different tasks and with short breaks in between. For this study infants were presented with videos of female faces displaying different communicative signals typically found in the infants' environment. Four trial types were presented to the infants twice (with each repetition being presented by a different actress). A fixation stimulus accompanied by attractor noises preceded trial presentation where the experimenter ensured that the infant was fixating at the centre of the screen. Each trial began with a 5-s baseline period where the face was still. The baseline was intended to draw the infants' attention to the screen and familiarize them with the face but was not included in the analysis. The baseline was followed by one of four dynamic sequences lasting ~16 s: (i) the eyes displayed gaze shifts towards or away from the infant while no other face part was moving, (ii) the mouth displayed vowel articulation movements while no other face part was moving, (iii) the hands placed next to the face displayed upward to downward motion while moving the fingers while no other face part was moving and (iv) the eyes, mouth and hands moved displaying a 'peek-a-boo' sequence. Pseudorandom presentation continued for a maximum of eight total trials of each sequence per infant.

Looking behaviour was recorded with a 17-inch flat-screen Tobii eye tracker. Gaze direction of each eye was measured separately, and from these measurements, the Tobii system evaluated where on the screen the individual was looking. During the task, the infant was seated on the parent's lap, at 50–55 cm from the screen, with height and distance of the screen adjusted to obtain good tracking of the eyes. A five-point calibration sequence was run, with recording and presentation of the study stimuli only starting when at least four calibration points were marked as properly attuned to each eye. Gaze data were recorded at 50 Hz. Fixations were defined automatically using temporal (100 ms) and spatial (35 pixels) filters. Clearview software was used for gaze data extraction. Areas of interest (AOIs) were defined around the eye, mouth and hands regions (covering the remaining non-face regions), and these were contrasted with another AOI covering all other areas of the face. Trials were excluded if <1 s of data was accumulated. Infants were excluded from the analysis if they were not administered the task or completed no valid trials. The majority of those included in the analysis completed the maximum number of trials (average trial count = 7.5) and accumulated 8–11 s of valid looking time data in each trial (see Supplementary materials).

Calculation and preliminary analysis of the eye–mouth index

To measure differences in looking to the eyes *vs* the mouth across the four conditions, an eye–mouth index (EMI) was calculated as follows: (looking time towards the eyes − looking time towards the mouth)/total looking time to any area of the screen. While it is well established that infants spend most of their time on internal features of the face, we scaled the measure by total looking time to ensure that any unusual behaviour in scanning of other features is accounted for using the same measure. The measure was derived for each trial and averaged across trials for each infant.

ANALYSIS AND RESULTS

Preliminary analysis was first conducted across the entire group to explore the extent to which the EMI measure was modulated by the different conditions across the two age groups. A general linear model included the repeated measures factors age (7 months *vs* 14 months) and condition (peek-a-boo, eyes, mouth and hands). After correcting for multiple comparisons, there was a significant interaction between age and condition [$F(2.9,\ 213.3) = 4.1$, $P < 0.001$]. When only the eyes were moving, infants spent 44% longer looking to the eyes relative to the mouth at 7 months. This amount increased slightly and non-significantly to 51% by 14 months. When the mouth was moving, infants spent 34% longer looking at the mouth relative to the eyes at 7 months, which rose significantly to 50% by 14 months [$F(1,78) = 7.9$, $P < 0.001$]. Across both ages, when a peripheral feature was moving (the hand condition) or when multiple features were moving (*peek-a-boo* scenes) infants preferentially look at the eyes (7 months: hands = 20%, peek-a-boo = 25%; 14 months: hands = 22%, peek-a-boo = 25%). This result confirms a general tendency to look more towards the eyes relative to the mouth across both age groups but shows that when only the mouth is moving, this general tendency is reversed where infants redirect their attention to the mouth. Estimated means for each group are shown in Figure 1 and suggest strong context modulation of EMI across all groups. The EMI values derived for each condition were used in subsequent analyses testing specific hypotheses.

We tested four inter-related hypotheses. The first was whether the amount of looking towards the eyes *vs* the mouth in complex dynamic scenes within the first year relates to risk group or to later outcomes. Second, we predicted that variability in looking towards the eyes relative to the mouth at both 7 and 14 months during a familiar and contextually rich *peek-a-boo* scene would predict 36-month EL in children, regardless of their clinical outcome. Previous studies have not investigated longitudinal change during the second year of life so we also examined, using the same paradigm, face scanning at 14 months of age. To assess the specificity of this prediction to EL, we included receptive language (RL) and controlled for non-verbal ability at 36 months, using a *t*-score derived from the Mullen Scales [Non-Verbal T-Score (NVT); see Supplementary material for details] at 36 months in each model. Third, a novel aspect of our study was to explore the origins of individual differences in infant scanning of faces. We compared, within individual, EMI in *peek-a-boo*, relative to simpler scenes where different face features are manipulated independently: moving eyes, moving mouth and moving hands. We expected that scanning of *peek-a-boo* would be a better predictor of language outcomes because it engages more endogenous control mechanisms than scanning simpler scenes. Fourth, we explored dimensional associations between face scanning in infancy and the degree of emerging autism symptoms as measured by the Autism Diagnostic Observation Schedule (ADOS) at 36 months of age.

Hypothesis 1. Eye vs mouth scanning in peek-a-boo predicts risk status or clinical diagnosis

A saturated path analysis model using a WLSMV estimator was used to examine the relationship between EMI and risk group membership, controlling for NVT at 36 months. Standardized model results showed NVT at 36 months to be a significant predictor of group [odds ratio (OR) = 0.79, 95% CI = 0.64−0.97, $P = 0.02$) but there was no significant relationship between group and EMI at 7 months (OR = 0.91, 95% CI = 0.51−1.63, $P = 0.76$) or 14 months (OR = 0.85,

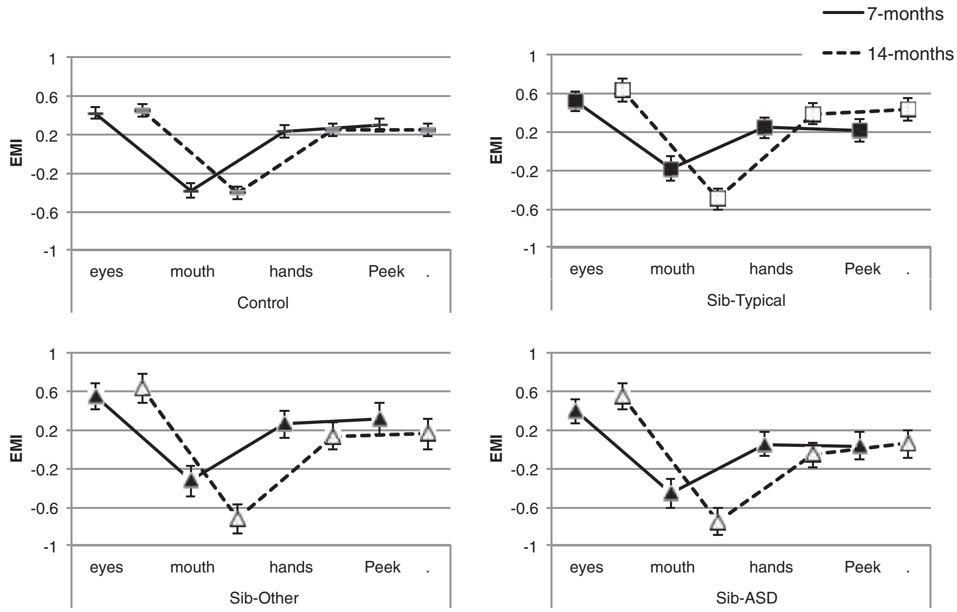

Fig. 1 Eye to mouth ratio (EMI) was derived as the relative looking time towards the eye *vs* the mouth scaled to the total amount of looking. EMI scores averaged over trials for each infant are shown. A score of +1.0 indicates 100% of eye–mouth time spent on the eyes, and a score of −1.0 indicates looking only to the mouth. Group differences in EMI were not significant.

95% CI = 0.45− 1.46, $P = 0.57$). EMI at 7 months did not predict EMI at 14 months (ß = 0.144, s.e. = 0.097, $P = 0.15$).

The relationship between EMI and 36-month outcomes was tested using a multinomial logistic regression model. Listwise deletion was used (because one high-risk child had no EMI score at 7 months or outcome data at 36 months) with a robust maximum likelihood estimator. The model showed no significant relationship between *peek-a-boo* EMI at 7 and 14 months and outcome group for any of the outcome contrasts (control *vs* Sib-ASD, Sib-TD and Sib-Other; all $P > 0.16$). Similar to previous studies, the null result held in the current dataset where we found no significant relationship between EMI and risk group or outcomes.

Hypothesis 2. Eye vs mouth scanning in peek-a-boo predicts EL

Hypothesis 2 attempts to replicate the previously reported findings by Young *et al.* (2009) that *peek-a-boo* EMI at 6 months predicted later expressive but not RL in at-risk and control groups. To test this hypothesis, an autoregressive cross-lagged path analysis model, with EMI at 7 and 14 months predicting 36-month RL and EL, controlling for outcome and 36-month NVT, was run using maximum likelihood estimation. The model was saturated, and standardized output showed *peek-a-boo* EMI at 7 months to be significantly associated with subsequent EL (ß = −2.47, s.e. = 0.08, $P = 0.01$), but not RL (ß = −1.09, s.e. = 0.09, $P = 0.27$). Negative *peek-a-boo* EMI (more looking towards the mouth relative to the eyes) at 7 months predicted superior EL at 36 months. In contrast, EMI at 14 months did not predict either EL (ß = −0.28, s.e. = 0.08, $P = 0.78$) or RL (ß = −0.66, s.e. = 0.09, $P = 0.51$).

Hypothesis 3. The relationship between EMI and language outcome is context-dependent

Peek-a-boo scenes are highly complex, encompassing several co-occurring signals on both the face and the hands and are therefore expected to require endogenous-orienting mechanisms. These scenes are also special in the infants' repertoire and are likely to reflect effects of social learning. As such, simpler manipulations of single face features, which are likely to rely less on endogenous control, could reveal the nature of the associations observed between variability in *peek-a-boo* EMI in infancy and later EL outcomes.

We tested the hypothesis that the relationship between EMI and later language outcome is context-dependent: complex communicative scenes requiring more endogenous control (i.e. peek-a-boo) differentially predict language outcomes relative to simple feature conditions (i.e. mouth, hand and eyes). A saturated path analysis model with a maximum likelihood estimator was used to examine the relationship between 7-month EMI in the *peek-a-boo*, mouth, eyes and hand conditions and 36-month EL and RL, controlling for NVT score and outcome. The relationship between 7-month *peek-a-boo* EMI and 36-month EL demonstrated in the previous model remained significant even after controlling for EMI in the single feature conditions (ß = −2.16, s.e. = 0.10, $P = 0.03$). More negative EMI (more looking to the mouth) in this complex condition was a better predictor of 36-month EL, but again no relationship with RL was found. Notwithstanding this pattern, the opposite association was observed for EMI in the mouth condition (ß = 2.44, s.e. = 0.09, $P = 0.02$): *more* scanning of the mouth when it alone was moving predicted worse subsequent EL. Here too, there was no significant association between mouth EMI and RL. EMI in the hand and eye conditions did not predict EL or RL. Clinical outcome group and NVT score did not significantly correlate with peek-a-boo EMI or with EMI in the single feature conditions.

Hypothesis 4. Context-dependent face scanning is associated with emerging autism symptoms

Finally, we tested whether face scanning was associated with the degree of social and communication skills as measured by the ADOS-G at 36 months. We excluded the control group from this analysis because the ADOS is a measure of clinical symptoms and may not be sensitive in the control group. Partial correlations between continuous 36-month ADOS-G total social communication score and EMI scores (in the four conditions; peek-a-boo, mouth, hand and eyes) were run, controlling for 36-month NVT. Notably, despite its association with 36-month EL, 7 months *peek-a-boo* EMI was not significantly associated with ADOS scores ($r = 0.09$, $P = 0.63$). In contrast, excessive scanning of the moving mouth in the mouth condition (which was associated with poor EL outcomes in the overall group) was also associated with a more severe emerging social and communication impairment measured by the ADOS ($r = -0.39$, $P = 0.03$) in the at-risk group. There was no association between scanning in the eyes or hands condition and ADOS scores.

DISCUSSION

It is widely accepted that the acquisition of communication skills in general, and language in particular, relies on the infants' ability to orient to relevant cues, ignore irrelevant ones and understand their referential nature. Typically developing infants successfully employ communicative signals to learn words from about 16 months of age (Baldwin, 1991, 1993). Infant's ability to follow gaze and engage in joint attention predicts later vocabulary size (Carpenter *et al.*, 1998; Morales *et al.*, 1998; Mundy *et al.*, 2000). Recent advances in theoretical modelling of brain and behavioural development coupled with advances in eye-tracking methodology have presented opportunities for understanding how infants selectively attend to various features of their complex environment to develop impressive social and communication skills. Our longitudinal study of infants from families with and without a family history of autism has offered new insights into how scanning of social scenes early on in infancy is associated with subsequent outcomes both in at-risk and low-risk infants.

Our group-level findings replicated and extended previous studies (Young *et al.*, 2009), suggesting that infants who go on to develop autism do not differ in their scanning of complex and dynamic social scenes such as *peek-a-boo*, nor in simpler scenes with single facial feature movements, such as the eyes, mouth or hand. Irrespective of their risk group or clinical outcome, infants exhibited clear modulation of their looking behaviour, i.e. looking to the eyes *vs* the mouth according to context, despite generally looking more towards the eyes in all contexts except when the mouth only is moving.

We further explored the origins of this relationship using context-dependent modulation of communicative signals in the same group of infants. The findings suggest that *peek-a-boo* EMI is strongly related to infants' scanning in single feature conditions, i.e. when only the eyes, mouth or hands are moving. We confirmed that the relationship between EMI and communication outcomes is context-dependent. Even after controlling for infants' EMI scores in simpler scenes, more looking to the mouth in *peek-a-boo* at 7 months still significantly predicted better EL. We took this pattern as supporting our hypothesis that the relationship between *peek-a-boo* EMI and EL is likely to be driven by the enhanced endogenous control required in more complex scenes. This pattern is similar to previous suggestions of the importance of cue integration, such as audio-visual cues measured using the McGurk effect (Kushnerenko *et al.*, 2008). While previous findings have specifically focused on the role of eye cues, such as gaze direction as precursors to language (Meltzoff and Brooks, 2008), our study highlights that endogenous orienting in complex scenes may be a more general precursor, at least as far as EL is concerned. The infants' greater endogenous control may enhance their ability to select relevant features and their ability to predict changes in the environment.

In contrast, those infants who were overly driven by exogenous factors such as mouth motion in single feature scenes exhibit poor EL, and within the at-risk group more pronounced symptoms of autism in toddlerhood. The latter findings are consistent with our previous studies using a non-social task with an independent group of infants at risk (Holmboe *et al.*, 2010). We observed subtle differences in the same independent at-risk group at 10 months of age, of which, preference for a repetitive central stimulus was predictive of greater social and communication impairment at 36 months (Elsabbagh *et al.*, 2011).

While causal links between looking behaviour in infancy and later childhood outcomes are tenuous, our study helps to reconcile paradoxical findings previously reported in literature on eye tracking and autism reviewed in the introduction. Our findings suggest that context sensitivity of scanning behaviour is influenced by individual variation in endogenous and exogenous orienting. On the one hand, more looking to the mouth in complex scenes that have multiple moving features and require a high degree of endogenous control was associated with superior language development across typical and atypical development. On the other hand, more looking to the mouth in simple scenes where the mouth is moving reflects stronger exogenous influence on scanning related to later development of poor EL across groups, and more specifically with severe social and communication impairment in childhood within the at-risk group as measured by the ADOS. It is important to note, however, that the relationship between exogenous orienting and outcomes did not hold equally across conditions or age groups. Unlike the mouth condition, we observed no such relationships in the hand and eyes conditions. Furthermore, 14-month EMI scores were unrelated to outcomes. This pattern of results reinforces the notion that attentional influence on developmental outcomes are most likely modulated by a combination of default biases and subsequent learning that modifies these biases over development. It is likely that within the early developmental period when language skills are emerging, mouth cues play a more important role relative to eye and hand cues (Lewkowicz and Hansen-Tift, 2012).

Our study, consistent with previous findings, suggests that scanning of complex social scenes does not begin as strikingly different in those infants later affected by autism (Young *et al.*, 2009). However, it is still possible that scanning of complex social scenes becomes increasingly different as a function of atypical interactions with the social environment over development. Supporting this pattern are recent findings suggesting atypical brain response to dynamic eye gaze in infancy, prior to the onset of autism symptoms (Elsabbagh *et al.*, 2012). Importantly, our current study highlighted the limitations of group analyses, which may often conceal important patterns of individual differences. In our study, the group of at-risk infants who developed ASD were highly variable in their EMI, a finding that could have been used to discount the relevance of these data to the development of infants at risk for autism. However, it is because the infants in both groups showed a wide range of variability in their looking behaviour as infants, as well as in their language outcomes, that we were able to capture clear associations between the two.

We replicated the observation that more looking towards the mouth relative to the eyes in dynamic communicative scenes predicts superior later EL (Young *et al.*, 2009). However, this is not a pattern specifically related to autism, nor reflecting compensatory strategies: The association between peek-a-boo EMI and later EL held across low-risk and at-risk groups and was not associated with the degree of social and communication impairment within the at-risk group.

Longer-term follow-up of our cohort may reveal further differentiation of the relationship between looking behaviour and the developmental trajectory of autism symptoms into later childhood. Our study raises additional questions that need to be addressed in future research. First, the association between endogenous control and language development was restricted to EL but was absent in RL. This finding converges with previous findings using a different testing environment where the infant was interacting with his/her caregiver (Young *et al.*, 2009). The reasons for this dissociation between EL and RL are unclear but may reflect finer variation in individual differences in expressive relative to RL, a hypothesis that needs to be verified in future studies. As such, different eye-tracking contexts or different at-risk populations may be needed to clarify this issue further. Second, it is not clear whether the observed association between the infant's eye-tracking behaviour and later language are specific to tracking in social scenes or if more general attentional abilities are also relevant. Finally, we only used total looking time but other measures of tracking, such as dwell time that require a finer resolution of data extraction procedures, may offer further insights into the different cortical processes underlying eye-tracking data.

SUPPLEMENTARY DATA

Supplementary data are available at *SCAN* online.

REFERENCES

Baldwin, D.A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, *62*(5), 875–90.

Baldwin, D.A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, *20*(2), 395–418.

Carpenter, M., Nagell, K., Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4), i–vi, 1–143.

Charman, T., Baron-Cohen, S., Swettenhamet, J., Baird, G., Drew, A., Cox, A. (2003). Predicting language outcome in infants with autism and pervasive developmental disorder. *International Journal of Language & Communication Disorders/Royal College of Speech & Language Therapists*, *38*(3), 265–85.

Chawarska, K., Shic, F. (2009). Looking but not seeing: atypical visual scanning and recognition of faces in 2 and 4-year-old children with autism spectrum disorder. *Journal of Autism and Developmental Disorders.* http://www.ncbi.nlm.nih.gov/pubmed/19590943 [accessed October 28, 2011].

Csibra, G., Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–53.

Dalton, K.M., Nacewicz, B.M., Johnstone, T., et al. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience*, *8*(4), 519–26.

Dalton, K.M., Nacewicz, B.M., Alexander, A.L., Davidson, R.J. (2007). Gaze-fixation, brain activation, and amygdala volume in unaffected siblings of individuals with autism. *Biological Psychiatry*, *61*(4), 512–20.

Dapretto, M., Davies, M.S., Pfeifer, J.H., et al. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, *9*(1), 28–30.

Deaner, R.O., Platt, M.L. (2003). Reflexive social attention in monkeys and humans. *Current Biology*, *13*(18), 1609–13.

Elsabbagh, M., Johnson, M.H. (2010). Getting answers from babies about autism. *Trends in Cognitive Sciences*, *14*(2), 81–7.

Elsabbagh, M., Volein, A., Holmboe, K., et al. (2009). Visual orienting in the early broader autism phenotype: disengagement and facilitation. *Journal of Child Psychology and Psychiatry*, *50*(5), 637–42.

Elsabbagh, M., Holmboe, K., Gliga, T., et al. (2011). Social and attention factors during infancy and the later emergence of autism characteristics. *Progress in Brain Research*, *189*, 195–207.

Elsabbagh, M., Gliga, T., Pickles, A., et al. (2013). The development of face orienting mechanisms in infants at-risk for autism. *Behavioral and Brain Research*, *251*, 147–54.

Elsabbagh, M., Mercure, E., Hudry, K., et al. (2012). Infant neural sensitivity to dynamic eye gaze is associated with later emerging autism. *Current Biology*, *22*(4), 338–42.

Falck-Ytter, T., Fernell, E., Gillberg, C., von Hofsten, C. (2010). Face scanning distinguishes social from communication impairments in autism. *Developmental Science*, *13*(6), 864–75.

Farroni, T., Menon, E., Johnson, M.H. (2006). Factors influencing newborns' preference for faces with eye contact. *Journal of Experimental Child Psychology*, *95*(4), 298–308.

Farroni, T., Csibra, G., Simion, F., Johnson, M.H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9602–5.

Gliga, T., Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science*, *20*(3), 347–53.

Hadjikhani, N., Snyder, J., Chabris, C.F., Clark, J., Steele, S., Vengel, M. (2004). Activation of the fusiform gyrus when individuals with autism spectrum disorder view faces. *NeuroImage*, *22*(3), 1141–50.

Hadjikhani, N., Joseph, R.M., Snyder, J., Flusberg, H.T. (2007). Abnormal activation of the social brain during face perception in autism. *Human Brain Mapping*, *28*(5), 441–9.

Haith, M.M., Bergman, T., Moore, M.J. (1977). Eye contact and face scanning in early infancy. *Science (New York, N.Y.)*, *198*(4319), 853–5.

Holmboe, K., Elsabbagh, M., Volein, A., et al. (2010). Frontal cortex functioning in the infant broader autism phenotype. *Infant Behavior & Development*, *33*(4), 482–91.

Hunnius, S., Geuze, R.H. (2004). Developmental changes in visual scanning of dynamic faces and abstract stimuli in infants: a longitudinal study. *Infancy*, *6*(2), 231–55.

Johnson, M.H. (1990). Cortical maturation and the development of visual attention in early infancy. *Journal of Cognitive Neuroscience*, *2*(2), 81–95.

Johnson, M.H., Griffin, R., Csibra, G., et al. (2005). The emergence of the social brain network: evidence from typical and atypical development. *Development and Psychopathology*, *17*(3), 599–619.

Joseph, R.M., Tanaka, J. (2003). Holistic and part-based face recognition in children with autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *44*(4), 529–42.

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. *Archives of General Psychiatry*, *59*(9), 809–16.

Kushnerenko, E., Teinonen, T., Volein, A., Csibra, G. (2008). Electrophysiological evidence of illusory audiovisual speech percept in human infants. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(32), 11442–5.

Langdell, T. (1978). Recognition of faces: an approach to the study of autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *19*(3), 255–68.

Lewkowicz, D.J., Hansen-Tift, A.M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences.* http://www.pnas.org/content/early/2012/01/13/1114783109 [accessed February 26, 2012].

Loveland, K.A., Landry, S.H. (1986). Joint attention and language in autism and developmental language delay. *Journal of Autism and Developmental Disorders*, *16*(3), 335–49.

Maurer, D., Salapatek, P. (1976). Developmental changes in the scanning of faces by young infants. *Child Development*, *47*(2), 523–7.

Meltzoff, A.N., Brooks, R. (2008). Self-experience as a mechanism for learning about others: a training study in social cognition. *Developmental Psychology*, *44*(5), 1257–65.

Merin, N., Young, G.S., Ozonoff, S., Roger, 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*(1), 108–21.

Morales, M., Mundy, P., Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development*, *21*(2), 373–7.

Mundy, P., Card, J., Fox, N. (2000). EEG correlates of the development of infant joint attention skills. *Developmental Psychobiology*, *36*(4), 325–38.

Neumann, D., Spezio, M.L., Piven, J., Adolphs, R. (2006). Looking you in the mouth: abnormal gaze in autism resulting from impaired top-down modulation of visual attention. *Social Cognitive and Affective Neuroscience*, *1*(3), 194–202.

Nummenmaa, L., Calder, A.J. (2009). Neural mechanisms of social attention. *Trends in Cognitive Sciences*, *13*(3), 135–43.

Pelphrey, K.A., Sasson, N.J., Reznick, J.S., Paul, G., Goldman, B.D., Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders*, *32*(4), 249–61.

Riby, D., Hancock, P.J.B. (2009). Looking at movies and cartoons: eye-tracking evidence from Williams syndrome and autism. *Journal of Intellectual Disability Research: JIDR*, *53*(2), 169–81.

Robertson, S.S., Guckenheimer, J., Masnick, A.M., Bacher, L.F. (2004). The dynamics of infant visual foraging. *Developmental Science*, *7*(2), 194–200.

Senju, A., Johnson, M.H. (2009). Atypical eye contact in autism: models, mechanisms and development. *Neuroscience and Biobehavioral Reviews*, *33*(8), 1204–14.

Speer, L.L., Cook, A.E., McMahon, W.M., Clark, E. (2007). Face processing in children with autism: effects of stimulus contents and type. *Autism: The International Journal of Research and Practice*, *11*(3), 265–75.

Spezio, M.L., Adolphs, R., Hurley, R.S., Piven, J. (2007). Analysis of face gaze in autism using "Bubbles". *Neuropsychologia*, *45*(1), 144–51.

van der Geest, J.N., Kemner, C., Verbaten, M.N., van Engeland, H. (2002). Gaze behavior of children with pervasive developmental disorder toward human faces: a fixation time study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *43*(5), 669–78.

Yarbus, A.L. (1967). *Eye Movements and Vision.* New York: Plenum.

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*(5), 798–814.