

“Are you looking at me?” How children’s gaze judgments improve with age.

Isabelle Mareschal¹, Yumiko Otsuka², Colin W.G. Clifford² & Denis Mareschal³

1 School of Biological and Chemical Sciences, Psychology, Queen Mary University of London, UK.

2 School of Psychology, UNSW Australia, NSW, Australia.

3 Centre for Brain and Cognitive Development, Department of Psychological Science, Birkbeck University of London, UK

Authors’ note: We are grateful to all the children and Melcombe Primary, Longfleet Primary, and Barnes Summer Play School who took part in this study. We thank Sinead Rocha and Hailey White for help with collecting the data. This work was funded in part by a Leverhulme project grant to IM, a Royal Society-Wolfson research merit award to DM and an Australian Research Council Discovery Project and Future Fellowship to CWGC.

Address all correspondence to Dr. Isabelle Mareschal, School of Biological and Chemical Sciences, Psychology, Queen Mary University of London, UK., Email: i.mareschal@qmul.ac.uk

Abstract

Adults' judgments of another person's gaze reflect both sensory (e.g. perceptual) and non-sensory (e.g. decisional) processes. We examined how children's performance on a gaze categorization task develops over time by varying uncertainty in the stimulus presented to 6- to 11- year olds (n=57). We found that younger children responded "direct" over a wider range of gaze deviations. We also found that increasing uncertainty led to an increase in "direct" responses, across all age groups. A simple model to account for these data revealed that although younger children had a noisier sensory representation of the stimulus, most developmental changes in gaze were due to a change in children's response criteria (category boundaries). These results suggest that although the core mechanisms for gaze processing are already in place by the age of 6, their development continues across the whole of childhood.

Being able to judge the direction of someone else's gaze plays a fundamental role in social interactions. It reveals where other people are looking and what is drawing their attention, and is therefore a strong social signal to their intentions and future actions (Baron-Cohen, 1995). Interestingly, the ability to detect and orient to direct gaze appears from a very early age. Infants as young as 5 months old spent longer looking at the eye region of smiling faces when they were making direct eye contact than when the eyes were averted (Symons et al. 1998). More recently, Farroni et al. (2002) reported that newborn babies also spent longer looking at faces when their gaze was direct than when it was averted. In addition, these authors also found that the ERP signals differed between direct and averted gaze in babies as young as 4 months of age (Farroni et al. 2004). Using a forced choice procedure, Doherty et al. (2009) reported that by 3 years of age children could reliably determine which of two faces was looking at them.

More recent methods for measuring how humans perceive another's gaze have led to the concept of a *cone of direct gaze* -- the range of gaze deviations that observers judged as being directed at them. It has been derived using either a gaze decentering/centering technique (e.g. Gamer & Hecht, 2007) or a gaze categorization technique (Ewbank et al. 2009; Stoyonova et al. 2010; Mareschal et al. 2013a). The categorization method has also been applied to study the development of gaze perception in children. For example, Vida and Maurer (2012) employed it to measure developmental changes in the ability to discriminate between direct and averted gaze along the horizontal (left-right) and vertical (up-down) dimensions. They reported that the horizontal cone of direct gaze was wider in children under the age of 6 than in

adults, suggesting a later development for fine-grained sensitivity to gaze. A binary categorization (left or right of forward) method has also been successfully used to measure perceived head direction in children (Vida et al. 2014).

Children's visual systems differ significantly from adults'. The majority of research points to the limiting role of internal noise on performance in infants (16 weeks and above) and young children (e.g. Skoczenski & Norcia, 1998; Beazley et al., 1980; Buss et al. 2008; Vida et al. 2014) rather than immaturities of the eyes (Brown et al., 2009). Abramov et al. (1984) used a battery of psychophysical tests to examine whether differences in performance for children (aged 5-8) compared to adults were purely due to sensory explanations or could be accounted for by other (non sensory) factors. They found that children were less sensitive than adults on most tasks (e.g. contrast sensitivity function, flicker fusion) but that this mainly reflected changes in response strategy. Specifically, children's attention levels drifted in near threshold tasks, leading to an increase in guessing rate. This is largely consistent with more recent reports that by roughly 3 years of age visual acuity is nearly adult like (Brown & Lindsey, 1998). Given that healthy (normal) social interactions in children might be compromised by a lack of normal visual experience (e.g. Jones & Klin, 2013), it is important to determine how accuracy on a visual task that underlies social interactions (e.g. gaze judgments) develops. In support of this idea, it has been shown that children in certain clinical groups spend less time looking at the eyes (e.g. autism (ASD), Pellicano et al. 2013). Although, a recent report from children raised with blind parents suggests no simple link between a lack of eye contact and autism, as these children (studied longitudinally from 6-47 months old) did not display autistic

behaviours and were not impaired on tasks that involved scanning the eye region of seeing adults (Senju et al. 2013).

When making a judgment of gaze direction, adults are influenced by a number of different features, including the emotion expressed by the face (Ewbank et al., 2009), its gender (O'Toole, et al. 1996; Slepian et al. 2011) and whether they hear it speak (e.g. Stoyanova et al. 2010). Perhaps the largest influence on gaze direction is the head orientation, as initially pointed out by Wollaston (1824) who showed that identical eyes in differently angled heads appear to gaze in different directions. Since then it has been found that head orientation and eye deviation interact, with head generally shifting the perception of gaze (Langton et al. 2004; Todorovic, 2006; Otsuka et al. 2014). Adults also demonstrate a “stare-in-the-crowd” effect; they are faster at detecting a direct gazing face within a crowd of averted gazing faces than an averted (rightwards) gaze within a crowd of direct and averted (leftwards) gazes (von Grunau and Anston, 1995). A similar effect has been demonstrated in children. For example, Senju et al. (2008) reported that children (normal and with autism) were faster at detecting a direct gaze (“stare in the crowd” effect) than averted gaze when eyes were presented alone (no facial information). However when shown within a laterally viewed (rotated) full face, only children with ASD did not show the stare in the crowd effect for direct gaze. This suggests that, at least for ASD, the contextual information of the face interacts with processing of the eye region indicative of information combination (between head rotation and eye direction). The general relationship between head direction and gaze is compounded by the fact that eyes can (and generally do) move independently of the head. As such, head direction is not always a reliable predictor of (congruent) gaze direction.

Given the relevance of gaze in social interactions, we sought to examine what limits children's performance under different conditions of task difficulty. We did this by altering features of the stimulus and measuring the children's judgments of gaze. We previously developed (Mareschal et al. 2013a) a simple psychophysical model that accounts for gaze categorization by three factors: (1) the width of the categorical boundaries (e.g. the categorical boundaries for "averted leftwards" and "averted rightwards"), (2) the peak (the midpoint between the category boundaries, akin to an observers' internal bias of what they judge as direct), and (3) the observers' uncertainty about the stimulus that reflects changes in both the internal noise (of the observer) as well as the external noise (imposed on the stimulus). Here we apply this model to the children's data to track changes in performance on three different gaze tasks as a function of age.

Specifically we sought to test the following hypotheses:

- 1) Younger children will have a larger cone of direct gaze than older children or adults. This hypothesis relates directly to performance in our "baseline" (noiseless) condition.
- 2) Since children use the head to detect direct gaze (Senju. 2008), removing configural information ("head" context) will decrease the CoD. This hypothesis relates directly to performance in our "eyes-only" condition.
- 3) Increasing uncertainty on the stimulus will lead to changes in response strategy. This hypothesis relates directly to performance in our "noise" condition.

Methods

Participants

A total of 57 children from a broad range of ethnic backgrounds in 3 mixed SES state supported primary schools took part in this study and were tested individually within their school classroom. There were 23 children in the “5.5 years old” group (14 girls; $M = 6.6$ years, $SD=3.8$ months), 18 children in the “7.5 years old” group (12 girls; $M = 8.3$ years, $SD=2.9$ months) and 16 in the “9.5 years old” group (6 girls; $M = 10.7$ years, $SD= 4.8$ months). All children had normal or corrected to normal vision.

Apparatus and stimuli

A Dell and a Lenovo laptop computers running MATLAB™ (MathWorks Ltd) were used for stimulus generation, experiment control and recording subjects’ responses. The programs controlling the experiment incorporated elements of the PsychToolbox (Brainard, 1997). Stimuli were displayed on a Thinkpad Edge laptop (1366*768 pixels, refresh rate: 60 Hz) driven by the computer’s built-in Intel HD Graphics 4000 card, or on a Dell Latitude E6500 (1280*800 pixels, 60Hz) driven by a NVIDIA Quadro NVS Graphics card. The displays were calibrated using a photometer and linearized using look-up tables in software. At the viewing distance of 57cm, one pixel subtended 1.5 arcmin for both computers.

The stimuli consisted of:

a) Face stimuli: Eight grey-scale adult faces (4 male and 4 female) with neutral expressions were created with Daz software (<http://www.daz3d.com/>). One of the female faces is shown in Figure 1a and b displaying direct gaze along with the two most extreme gaze deviations tested ($\pm 20^\circ$). The hair was cropped and the face was

presented within a circular aperture in the middle of the monitor. The stimuli subtended on average $15.1 \text{ deg} * 11.2 \text{ deg}$ and were viewed at approximately 57 cm. Stimuli were uploaded into FaceGen and the (pixel) position of the iris for each forward gazing head was determined using Gimp software. In order to control the direction of gaze, the original eyes in the faces were replaced by grey-scale eye stimuli created using Matlab, making sure that the pixel location of the new iris was the same as that recorded in Gimp. The inter-ocular distance was kept the same as the original face and a small amount of vergence was added so that the left and right eyes converged at a distance of 57cm. The deviation of each eye was independently controlled using Matlab procedures that gave us precision down to the nearest pixel for eye rotation along the horizontal axis.

b) Noisy faces: Fractal noise ($1/f$ amplitude spectrum) was added to the eyes of the same faces (Figure 1b, middle). Since contrast sensitivity is dependent on age (e.g. Brown & Lindsey, 2009 for a review; Beazley et al. 1980) we tailored the strength of the noise to each child by changing the pupil/sclera contrast to ensure equal sensitivity across age groups (see *noise contrast task* below).

c) Eyes-only stimuli: In order to examine configural head influence, we presented only the (noiseless) eyes of the same eight faces in this condition (Figure 1b, right). The stimuli subtended on average $1.3 \text{ deg} * 7.2 \text{ deg}$ (the same size as when within the head context).

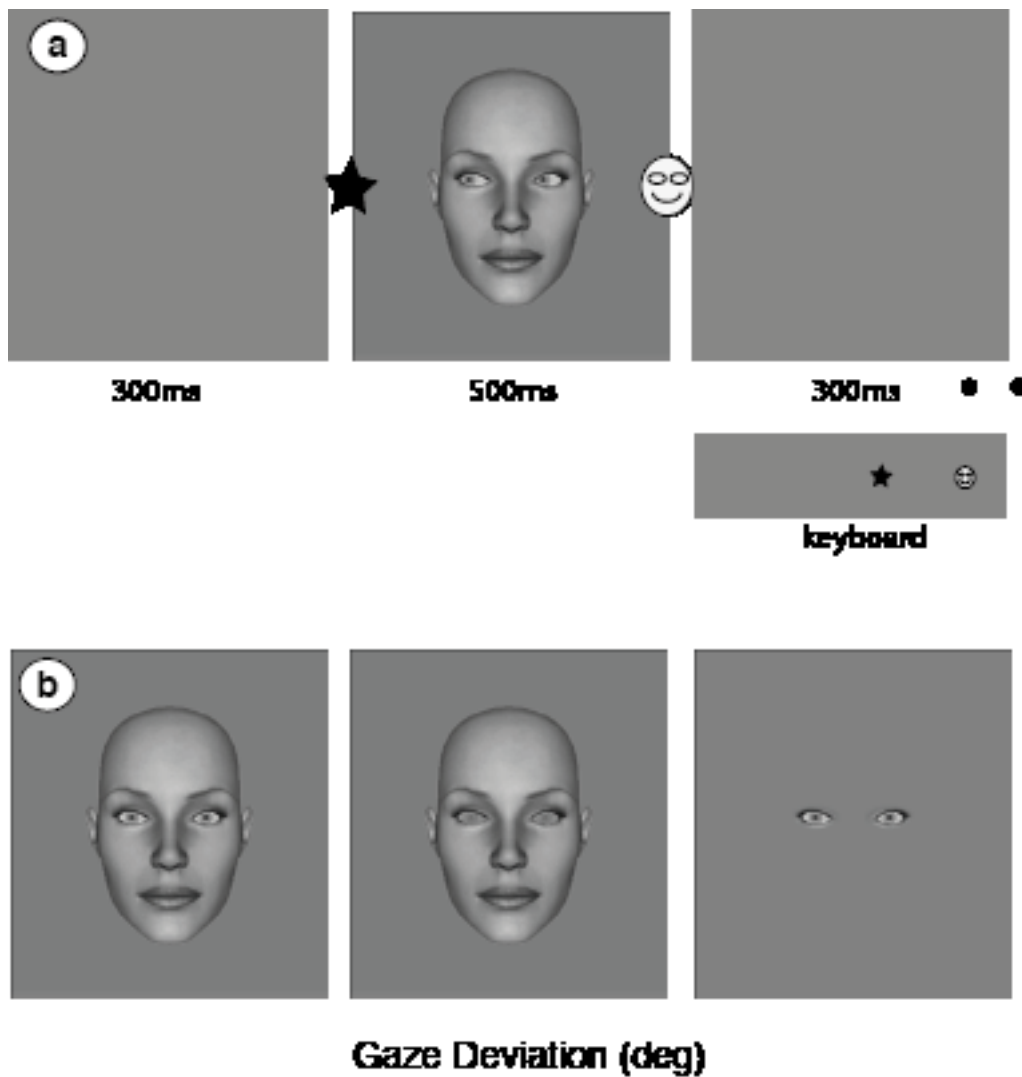


Figure 1. Experimental procedure. (a) In the main condition, a face appeared for 500ms followed by a grey screen for 300ms. At the end of the 300ms, the child could respond using a key press to indicate the direction of gaze. (b) Sample face in the noiseless (left), noisy condition (middle), and eyes-only (right). The Noisy condition was created by adding fractal noise to the eyes (shown here for the same female face). The Eyes-only condition was achieved by applying an elliptical raised cosine contrast envelope over each eye (same female face as in a).

Procedure

In order to compensate for differences in susceptibility to noise as a function of age we tailored the strength of the noise to each child, as described below.

a) Noise contrast task

Noise was added to the eyes of computer-generated faces (RMS contrast of 6%) and the children were asked to judge whether the gaze was to their left or to their right. One of four possible faces appeared with a fixed gaze deviation of either $+12^\circ$ (right) or -12° (left) for 500ms followed by a grey screen (300ms) after which the child could respond, giving their answer using a key-press (or verbally if they preferred to do so). We employed the Psi Bayesian adaptive procedure that estimates the choice of the next trial stimulus level based on the responses to all the previous trials. By optimizing the stimulus placement, the staircase is very efficient while being more robust to changes in slope and therefore is well suited to test young children. It converges at the end of 32 trials on the pupil/sclera contrast level leading to 80.3% correct discrimination (Konstevich and Tyler, 1999). The pupil/sclera contrast obtained by the staircase was subsequently used for that child's categorization tasks. This level could not exceed 0.4 and most children's thresholds were below this value (see Table 1 for average values).

b) Categorization task

The child's task was to indicate whether the direction of gaze in the three different conditions was averted to the left, direct, or averted to the right using 3 different key-presses. Each stimulus was presented for 500ms followed by a grey screen that lasted 300ms during which no response was recorded. There was no time pressure and the next trial was only initiated after a response was made. A pause was introduced after

10 trials, the screen was set to grey and the child was asked if they wanted to continue. Stimuli were presented using a method of constant stimuli with 9 different directions of gaze selected from the set: $\{-20^\circ, -9^\circ, -6^\circ, -3^\circ, 0^\circ, 3^\circ, 6^\circ, 9^\circ, 20^\circ\}$. Each direction of gaze was tested 8 times per condition. Not all children performed all conditions. To ensure continued engagement with the task, children in the 5.5 years old and 7.5 years old groups received stickers between testing conditions. Children in 9.5 years old group were offered stickers at the end of the session. Two stickers were placed on the edges of the monitor (Fig 1a), one on the left side (star) and one on the right side (smiley face). Two identical stickers were also placed on the two corresponding response keys. The child's task was to indicate the direction of gaze (towards the star; press star / towards the smiley face; press smiley face). Although the faces and stickers were in the same depth plane, the children had no difficulty indicating when gaze was directed towards one (or the other) sticker.

c) Instructions

Noise contrast task: "On the screen you're going to see some faces. Now, these faces aren't always going to be normal faces like yours or mine. They're going to have really funny looking eyes! Their eyes are going to be a bit fuzzy looking. Your job is to work out which way the eyes are looking. Sometimes it will be quite easy but other times it will be really hard, so concentrate really hard, but if you're not sure, you can just guess! Now, if you think the eyes are looking this way (points), towards the star, you press the star button. And if you think that the eyes are looking this way (points), towards the smiley face, you press the smiley face button. If they're looking this way, which button do you press? (*child responds*), well done! And if they are looking this way, which button do you press? (*child responds* - corrected if wrong and asked until

they get it right). Great! Let's start playing.”

Categorization task: “Well done, you did a really great job! Now, we're going to play again. If you think the eyes are looking this way (points) towards the star, you press the star button. And if you think that the eyes are looking this way (points), towards the smiley face, you press the smiley face button. Just like before. Except, this time, if you think that the eyes are looking straight at you, you can press the middle button, the K button (show button - buttons in use were “j”, “k” and “l”). So, that way (points) you press star, that way (points) smiley face, and straight at you, the middle button”. (Checks again that they understand).

Data analysis strategy

a) Logistic fit

We fit each child’s data using the (model free) conventional method (e.g. Stoyanova et al., 2010) where two logistic functions were fitted to the proportion of “left” and “right” responses. A function for “direct” responses was calculated by subtracting the sum of the “left” and “right” responses from 1.0 (Fig. 2b). It is important that these three functions be fitted as an ensemble (here using the Nelder-Mead simplex method (Nelder & Mead, 1965) implemented via Matlab’s *fminsearch* function, as in Mareschal et al., 2013a) to minimize residual variance and avoid introducing bias. The cross-over points of the “direct” and the “left” (L1), and “direct” and “right” (R1) responses respectively are termed the categorical boundaries and the separation between the two is taken as the cone of direct gaze.

b) Psychophysical Model:

We also fit each child's data using a psychophysical model that accounts for data based on three free parameters (Mareschal et al. 2013a). The advantages of this model are that it formalizes the assumptions that underlie conventional curve fitting and that all the data (left/right averted and direct) are fit together.

- 1) An estimate of the peak that corresponds to the midpoint between the category boundaries and represents the gaze direction most judged to be direct (also known as a person's bias).
- 2) An estimate of the width of direct judgments that corresponds to the distance between the categorical boundaries between leftwards and direct and rightwards and direct.
- 3) An estimate of the standard deviation of the observers' sensory representation of a gaze stimulus. This represents the uncertainty associated with the estimate and reflects the noise (internal and external) affecting the observer's sensory representation.

We note that in this type of task, any trial-by-trial shifts in the children's criterion (their category boundaries) may affect their responses and could cause an increase in the standard deviation (SD). As such our SD component encompasses both these factors in its estimate.

Results

Categorization Response Curves: logistic fits

We were not always able to collect a full set of data on all of the children. For the 5.5 years old group, 4 children were excluded resulting in usable data from 19 out of 23 (three children only performed the detection task, and one child had data that could not be fit (e.g. mainly responding in a single direction), 17 out of 18 for the 7.5 years old group (1 child stopped after the detection task), and 16 out of 16 for 9.5 years old group.

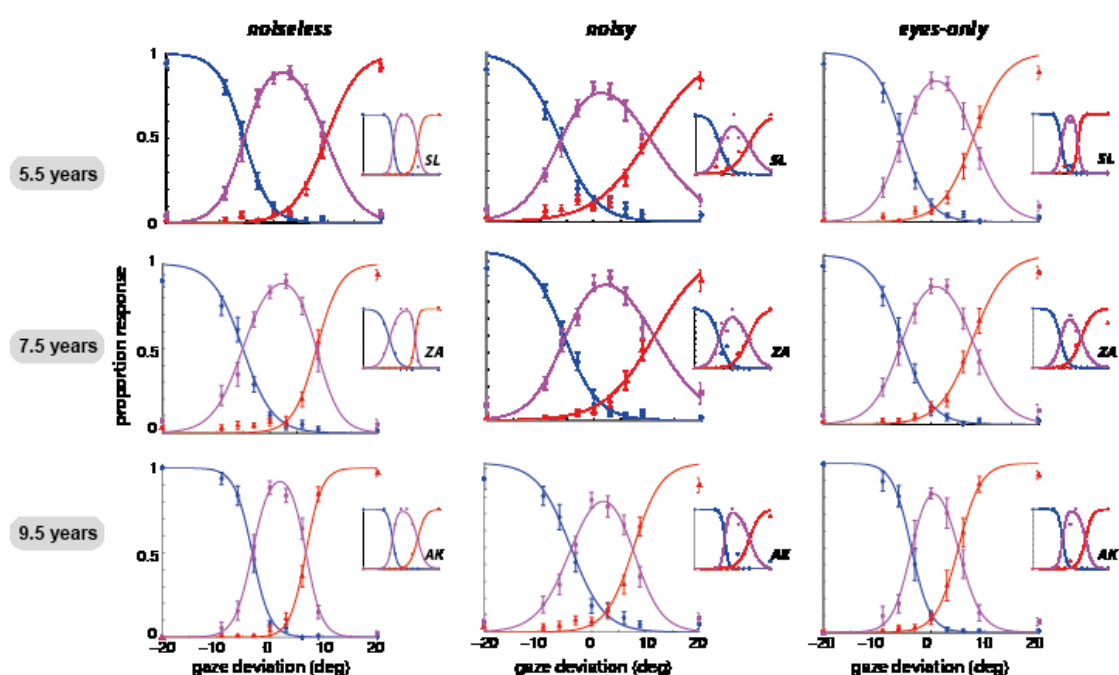


Figure 2 Categorization data for the Noiseless, Noisy and Eyes-only conditions for the three age groups. Data points are the proportion of responses to different directions of gaze: leftwards (blue diamonds), direct (pink squares) and rightwards (red triangles). (a) 5.5 years old. Large graphs are averaged data, smaller insets are data for one representative child (SL) in the three conditions. Error bars are ± 1 s.e.m. (b) 7.5 years old, inset data from ZA and (c) 9.5 years old, inset data from AK.

Figure 2 plots the data, averaged across all children within each age group, for the three different conditions. Insets are sample data from children who performed all three conditions in different age groups. Blue and red curves are logistic fits to the leftwards and rightwards responses respectively and purple is the direct responses. For each child we report the “cone of direct gaze” (CoD) that corresponds to the distance in degrees between the left and direct cross-over points (blue and purple curves) and right and direct cross-over points (red and purple curves). In order to examine changes in the overall number of direct responses, we also report the area under the direct response curve (AUC) for all children. Both are estimates of how many gaze deviations the child judged to be directed at them. Average values are given in Table 1 as well as the average noise contrast used. We note that the sclera/pupil contrast value used for 9.5 years old children of 0.24 is comparable to that used in adults that ranged between 0.17 and 0.20 (Mareschal et al. 2013a,b). Since the Psi-procedure is based on 32 trials, we also calculated the mean value of the standard error (of the noise threshold measure) for the three different age groups. These are (0.027) for 5.5 years old children, (0.043) for 7.5 years old children, and (0.049) for 9.5 years old children. In the context of the between subjects errors in thresholds reported in Table 1, these values indicate that only around 40% of that variance is due to measurement error.

In order to ensure that children were paying attention to the task, we also calculated the response rates for each condition and age group. This was taken as the average number of correct identifications in the most extreme (i.e. the two easiest) gaze deviation conditions. For difficult discrimination tasks in adults (such as visual crowding) these values are around 95-97% and can result from finger press errors, lapses in attention, or confusion about the target (Mareschal et al. 2010). We note that

in the Noiseless and Eyes-only condition, the response rates were above 90% for all age groups, with the 7.5 and 9.5 years old children performing close to adult levels. The response rates are lower in the Noisy condition, but this largely reflects the fact that the noise level was chosen to maintain 80.3% correct identification for a 12° deviation. The values here are slightly higher (since the extreme deviation was 20° degrees) but close to 80.3% suggesting that children were not attending less well or making more finger press errors in this condition than the other ones. As expected, with a fixed level of noise, the threshold pupil/sclera contrast (and hence signal-to-noise ratio) decreases with age.

Table 1: Cone of Direct Gaze (CoD) and AUC for all three conditions as a function of age

	Noiseless	Noisy	Eyes-only	Pupil/sclera contrast
5.5 y.o (CoD)	15.1±6.4	16.2±5.2	12.7±7.3	0.35±0.05
(AUC)	15.6±5.1	15.1±3.4	13.9±5.8	
Resp. Rate	92.3%	83.8%	90.4%	
7.5 y.o (CoD)	13.1±3.1	14.7±2.8	13.2±5.3	0.33±0.06
(AUC)	13.2±3.2	14.7±2.6	13.1±4.6	
Resp. Rate	91.5%	86.0%	91.5%	
9.5 y.o (CoD)	10.0±2.6	10.9±3.4	9.7±3.4	0.24±0.08
(AUC)	10.2±2.6	11.2±3.1	9.8±3.4	
Resp. Rate	97.6%	91.1%	94.7%	

Note: CoD is measured in degrees, averaged across all children who performed all three conditions. Response rates are the average percentage of correct gaze categorizations at the extremes ($\pm 20^\circ$).

Analyses using AUC or CoD yielded similar results, so we report the analysis for CoD only as this is the more commonly used measure. The data were analyzed using a mixed-design ANOVA with Condition as a repeated (within-subject) factor and Age-group as a between-subject factor. We found main effects of Condition ($F(2, 86) = 4.84, p < .01, \eta_p^2 = .10$) and of Age-group ($F(2, 43) = 4.18, p < .025, \eta_p^2 = .16$), with no significant interaction. Post-hoc Tukey tests revealed that the 5.5 year old children have significantly broader CoDs than 9.5 years old children ($p = .02$). CoDs for 5.5 and 7.5 years old did not differ significantly ($p = .78$), while CoDs for 7.5 and 9.5 years old children were marginally different ($p = .09$). Paired t-tests to explore the main effect of Condition revealed that CoDs for Eyes-only were significantly narrower than those for the Noisy condition ($t(45) = 3.54, p < .001$) and that CoDs for the Noiseless condition were significantly narrower than the CoDs for the Noisy condition ($t(47) = -1.9, p < .06$), but that CoDs for the Noiseless and Eyes-only conditions were not significantly different. Note that differences in degrees of freedom between t-tests occurred because children who only performed two of the three conditions could be included in the relevant paired t-test comparison for which they had provided data. The same applies to the paired t-test comparisons reported further in the results section.

Model fits

In order to examine what changes underlie children's performance as a function of condition at different ages, we fit the psychophysical model to each child's data. Figure 3 shows the model fit to sample data for two 5.5 years old children (top row), two 7.5 years old children (middle) and two 9.5 years old children (bottom row). The data were chosen to highlight the variability between children's data within and across age groups. Symbols are the children's responses to the different gaze deviations and curves are the model fits to the data in the Noiseless condition (left column), in the Noisy condition (middle) and in the Eyes-only (right column). Notice that there is more inter-observer variability for the younger ages than the older ages (e.g. AK and SR responses look more similar than SA and SL across the different conditions tested). Estimates of goodness of fit were calculated for the model fit to each child's data in each condition and the averages (of these individual fits) were calculated. For the 5.5 years old children, the model accounted for 92.6% of the variance (Noiseless), 90.6% (Noisy) and 93.9% (Eyes-only). For the 7.5 years old the model accounted for 90.2% (Noiseless); 81.7% (Noisy) and 88.5% (Eyes-only), for the 9.5 years old it accounted for 93.9% (Noiseless), 78% (Noisy) and 94.4% (Eyes-only). All children whose data could be fit with the model were included in the analysis.

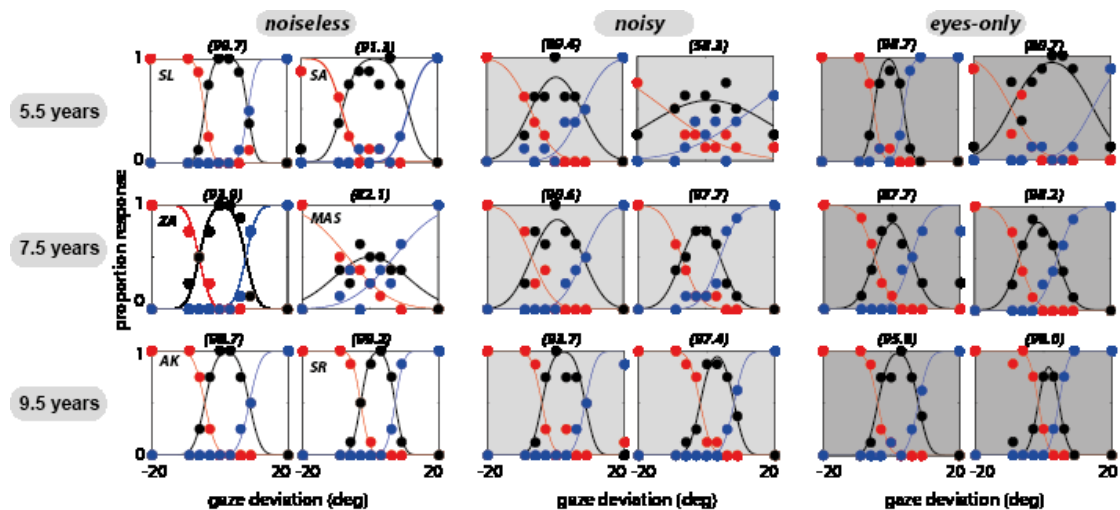


Figure 3 Model fits to two participants for each of the three age groups across the three conditions. The two examples per age group were chosen to show the most different types of responses to the three conditions. Values above plots are the percentage of variance accounted for by the model.

The model returned three parameters per child, per condition: (1) an estimate of the Peak (Bias), (2) the distance between the Category Boundaries (Width), and (3) the standard deviation (SD) of the sensory representation of gaze. In order to examine how the children’s data compares to adults’, we have also included previously reported parameter estimates obtained on an adult group using identical stimuli (mean age = 31.2 years; SD = 7.6 years; data taken from Mareschal et al. 2013a). Data and statistical analyses for the model parameters are shown only for children that completed all three conditions, which corresponds to 16 of the 5.5 years old children (70% of total), 17 for the 7.5 years old (94% of total) and 14 for the 9.5 years old (88% of total) totaling 47 children. In addition, in a very small number of cases (3) where one of the 3-parameter estimates returned an impossible value (either width of category boundaries greater than the range tested, or an SD of 0), all three parameters for the corresponding condition only were excluded. Estimates (averaged across participants’ fits) are presented in Table 2.

Table 2: Average peak, width and SD estimates in degrees

	Peak (noiseless)	Peak (noisy)	Peak (eyes-only)
5.5 y.o	2.16 ± 1.68	1.72 ± 1.14	1.28 ± 1.60
7.5 y.o	1.68 ± 1.35	2.72 ± 2.34	0.87 ± 1.68

9.5 y.o	1.50 ± 1.41	2.12 ± 2.32	0.60 ± 1.79
Adult	1.11 ± 0.89	1.37 ± 1.26	0.46 ± 0.81
	Width(noiseless)	Width (noisy)	Width(eyes-only)
5.5 y.o	15.65 ± 5.65	16.72 ± 5.11	13.78 ± 6.55
7.5 y.o	15.33 ± 6.30	16.69 ± 6.01	13.80 ± 5.22
9.5 y.o	10.01 ± 2.85	13.62 ± 6.84	9.24 ± 3.69
Adult	6.33 ± 2.05	6.93 ± 2.73	4.73 ± 1.28
	SD (noiseless)	SD (noisy)	SD (eyes-only)
5.5 y.o	3.69 ± 2.34	6.30 ± 4.06	3.89 ± 2.12
7.5 y.o	4.24 ± 3.03	5.26 ± 3.06	4.31 ± 3.08
9.5 y.o	2.16 ± 1.76	5.71 ± 6.17	2.35 ± 1.33
Adult	2.41 ± 0.66	4.40 ± 1.28	2.22 ± 0.44

We performed a mixed-design ANOVA with Condition as a repeated (within-subject) factor and Age-Group as a between-subject factor, on each of the 3 parameters of interest. The ANOVA revealed a significant effect of Condition for the Peak ($F(2, 98)=9.84, p<.001, \eta_p^2=.17$), Width ($F(2, 98)=15.62, p<.001, \eta_p^2=.24$), and SD ($F(2, 98)= 15.10, p<.001, \eta_p^2=.24$) parameters. There was also a main effect of Age group for the Width parameter only ($F(3, 49)=11.10, p=.001, \eta_p^2=.40$), and no interaction for any of the parameters.

Follow up posthoc t-tests (Tukey) revealed that 9.5 years olds Width was marginally significantly narrower than 5.5. years old ($p=.06$) and that adult widths were significantly narrower than for all other age groups ($p<.05$). T-tests on the Width of the Category Boundary revealed that all conditions differed significantly from each

other, with Eyes-Only being the narrowest condition and Noisy the widest (Noiseless significantly narrower than Noisy, $t(53)=-3.15, p<.003$; Eyes-only significantly narrower than Noisy, $t(52)=4.87, p<.001$ and significantly narrower than Noiseless $t(53)= 3.72, p<.001$). For the Peak values, t-tests revealed that the Noisy and Eyes-only conditions differed significantly ($t(52)=4.98, p<.001$) and the Noiseless and Eyes-only conditions ($t(54)=3.39, p<.001$). Finally t-tests on the SD values revealed that Noisy was significantly wider than Noiseless ($t(53)=-4.12, p<.001$) and also significantly wider than the Eyes-only condition ($t(52)=4.32, p<.001$).

Discussion

We report the following findings: (1) the cone of direct gaze narrows with age, consistent with Vida & Maurer using noiseless stimuli (2012); (2) increasing uncertainty by adding noise to the stimulus widens the cone of direct gaze; and (3) removing configural information (the head cue) narrows the cone of direct gaze. The latter two findings reveal that, when tested with identical stimuli, the children we examined performed similarly to adults (Mareschal et al. 2013a). When examined with the model, we report that the only significant effect of age is a narrowing of the Width parameter suggesting that children's response criteria and, by consequence, their category boundaries are what underlie the changes in the Cone of Direct gaze, not their sensory representation (e.g. their internal noise).

Using a categorization task, we report that the CoD is wider in young children and narrows with age. This is consistent with Vida & Maurer (2012), although they report narrower CoDs than here, possibly reflecting a difference in the two procedures (they

gave feedback during practice which may have led children to use a more stringent criterion). When noise is added to the stimulus, children increase their direct responses, as evidenced by a broadening of the CoD and a greater area under the curve measure. We have previously shown in adults that when they are uncertain about the stimulus, they have a prior expectation to assume that gaze is directed at them (Mareschal et al. 2013b; Mareschal et al, 2014). The result here can be largely accounted for by the effect of a widening of the category boundaries (Width) and an increase in the standard deviation of their representation of the stimulus (their uncertainty). There is growing evidence for a special status for direct gaze that is already present in babies (Farroni et al., 2002; Samuels, 1985; Vecera & Johnson, 1995). Interestingly, we also find that the peak (bias) of direct responses is positive (rightwards) in all conditions. Using photographs of real faces, Calder et al. (2008) have reported a small but consistent bias in adults in the same direction. They found that a 5° rightwards gaze deviation was more likely than a 5° leftwards one to be categorized as direct, and that a physically direct gaze was more likely to be categorized as leftwards than rightwards. Importantly, they also mirror flipped their images and the bias remained, suggesting that it is not simply the result of small asymmetries present in both real and avatar faces. It appears that a similar bias is evident in children's perception of gaze although it remains unclear what functional purpose this may serve.

When configural information (a forward facing head) is present, we find evidence that children use information about the head direction to judge the eye deviation, even for the 5.5. year old children. This is consistent with the finding that head orientation plays an important role in adults' perception of gaze (Mareschal et al, 2013a),

although the effect of the presence versus absence of a direct head on gaze judgments is not always evident (Otsuka et al. 2014). We also find, as with adults, that although the width of the category boundaries narrows in the Eyes-Only condition there is no concomitant increase in the SD, which might be expected since removing head information could increase uncertainty. However this depends on (a) the uncertainty associated with the head orientation judgment as well as (b) the amount of noise associated with the combination of head orientation and gaze deviation. Interestingly, in the children we tested, the use of the (forward) configural head cue appears to be reliable, consistent with previous findings that children generally rely strongly on the head direction information than eye deviation. Using a pointing and looking task, Doherty et al. (2009) found that head rotation affects childrens' judgments of where an adult is looking. They found that children as young as 3 performed better on the task when the head and gaze of the adult demonstrator were congruent. Evidence of children using head orientation is also present in Senju et al. (2008), who used a "stare-in-the-crowd" task to show that, for normally developing children, inversion effects abolish the advantage of detecting a direct gaze only when full faces are shown. When the eyes are shown alone the inversion effects disappear. This indicates that, in normally developing children, faces are processed holistically and that information about the head direction is used in their judgments of gaze. It is worth noting that a possible effect on our results is that when the heads are removed the Eyes-Only stimuli lack realism and that this modulates the effect. Although this is possible, recent work by Takahashi & Watanabe (2013) suggests that children orient to gaze from pareidolia faces (objects that look like faces) in a similar manner to cartoon and real faces.

We find a significant effect of age on the CoD, such that the cone of direct gaze narrows with age. Examination of the model parameters reveals a significant effect of age on the category boundaries only, with an average width for 9.5 years old children of 10° for noiseless faces, which is only slightly larger than adults (approximately 8°), suggesting that the narrowing is due to a change in response criterion rather than a change in children's sensory representation. Although we might have expected to find a significant effect of age on the SD, it is worth noting that we equated performance for the noise condition by changing the pupil/sclera contrast accordingly. The fact that this value was lower for the older children (i.e., the threshold signal-to-noise ratio was lower) than the younger ones indicates that the child's uncertainty due to noise in the sensory representation reduces with age.

In adults, it is believed that gaze direction is coded by a multi-channel system with at least 2 channels representing averted (leftwards and rightwards) directions of gaze and one explicitly representing direct gaze (Calder, Jenkins, Cassel and Clifford, 2008). In this framework, the perceived direction of gaze of, for example, a slightly rightwards gazing stimulus is determined by the relative activity of the three channels (near baseline in the leftwards channel, higher in the direct and right channels). An increase in uncertainty (e.g. noise) would lead to a slight increase in activity in the rightwards and direct channels with the same (baseline) amount of activity in the leftwards channel such that their *relative* activity is different to the noiseless condition (see Clifford et al. 2015 for a fuller discussion). This results in a shift towards the central tendency (e.g. "direct"); the observer classifies gaze deviations as "direct" over a larger range of gaze deviations (larger Width parameter). We note that the change in category boundaries need not be a "higher-level" (cognitive) effect. We

find that younger children (5.5. year olds) have a wider CoD, mainly due to an increase in their category boundaries. This in could result from a wider “direct” gaze channel, and/or greater weighting of the (forward) head direction cue.

One question that arises from this procedure is whether gaze judgments change when children are tested with other children’s faces, rather than adult Caucasian faces. For example, adults and children display an “other-race effect” (poorer at making judgments in other races), which can affect how they look at faces (e.g. Kelly et al. 2007; Fu et al. 2012; Suhrke et al. 2014). A similar effect has also been shown with age; faces similar in age to ones own receive more attention than those of a dissimilar age (Ebner at al. 2013; Rhodes & Anastasi, 2012). How these factors may influence children’s overall judgments of gaze remain to be investigated.

Overall our data suggest that the foundations for gaze judgments are present by the age of 6. However, given that we find evidence of developmental changes (the CoD narrows with age), it is likely that these processes get tuned, possibly via normal social interactions., For example, it has been reported that after an initial normal period, babies (ages 2-6 months) that develop autism spectrum symptoms (ASD) at a later stage, spend less time looking at the eyes faces (Jones & Klin, 2013). This suggests that abnormal sensory input may have knock on effects for the development of social interactions, although this interpretation remains speculative for the moment. Indeed, whether healthy social interactions inform the development of normal gaze behavior or the other way around, and whether there is an optimal period in time for these to take place remain to be determined.

References

Abramov, I., Hainline, L., Turkel, J., Lemerise, E., Smith, H., Gordon, J., & Petry, S. (1984). Rocket-ship psychophysics. *Investigative Ophthalmology & Visual Science*, 1307-1315.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA, MIT Press.

Beazley, L.D., Illingworth, D.J., Jahn, A. & Greer, D.V. (1980). Contrast sensitivity in children and adults. *British Journal of Ophthalmology*, 64, 863-866.
doi:10.1136/bjo.64.11.863

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433-436.

Brown, A.M. & Lindsey, D.T. (2009). Contrast insensitivity: the critical immaturity in infant visual performance. *Optometry Visual Science*, 86, 572-576. doi:
10.1097/OPX.0b013e3181a72980

Buss, E., Hall, J.W., & Grose, J.H. (2009). Psychometric functions for pure tone intensity discrimination: slope differences in school-aged children and adults. *Journal of Acoustic Society of America*, 125, 1050-1058. doi: 10.1121/1.3050273.

Calder, A. J., Jenkins, R., Cassel, A., & Clifford, C. W. G. (2008). Visual representation of eye gaze is coded by a non-opponent multichannel system. *Journal of Experimental Psychology: General*, 137(2), 244–261. doi: 10.1037/0096-3445.

Campbell, R., Lawrence, K., Mandy, W., Mitra, C., Jeyakuma, L., & Skuse, D. (2006). Meanings in motion and faces: Developmental associations between the processing of intention from geometrical animations and gaze detection accuracy. *Development and Psychopathology*, 18, 99-118. doi: 10.1017/S0954579406060068

Clifford, C.W.G., Mareschal, I., Otsuka, Y., & Watson, T.L. (2015). A Bayesian approach to person perception. *Consciousness and Cognition*, in press. doi: 10.1016/j.concog.2015.03.015

Doherty, M.J., Anderson, J.R. & Howieson, L. (2009). The rapid development of explicit gaze judgment ability at 3 years. *Journal of Experimental Child Psychology*, 104, 296-312. doi:10.1016/j.jecp.2009.06.004

Ebner, N.C., Johnson, M.R., Rieckmann, A., Durbin, K.A., Johnson, M.K., & Fischer, H. (2013). Processing own-age vs other-age faces: neuro-behavioral correlates and effects of emotion. *Neuroimage*, 78, 363-371. doi: 10.1016/j.neuroimage.2013.04.029

Ewbank MP, Jennings C, Calder AJ (2009) Why are you angry with me? Facial expressions of threat influence perception of gaze direction. *Journal of Vision*, 9, 1-7. doi: 10.1167/9.12.16.

Farroni, T., Csibra, G., Simion, F. & Johnson, M.H. (2002). Eye contact detection in humans from birth. *Proceedings of the National academy of Sciences USA*, 99, 9602-9605. doi: 10.1073/pnas.152159999

Farroni, T., Johnson, M.H. & Csibra, G. (2004). Mechanisms of eye gaze perception during infancy. *Journal of Cognitive Neuroscience*, 16, 1320-1326. doi:10.1162/0898929042304787

Fu, G., Hu, C.S., Wang, Q., Quinn, P.C., & Lee, K. (2012). Adults scan own-and other-race faces differently. *Plos One*, 7(6). doi: 10.1371/journal.pone.0037688

Gamer, M. & Hecht, H. (2007). Are you looking at me? Measuring the cone of gaze. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 705-715. doi: 10.1037/0096-1523.33.3.705

Gori, M., Del Viva, M., Sandini, G., & Burr, D. (2008). Young children do not integrate visual and haptic form information. *Current Biology*, 18, 694-698. doi: 10.1016/j.cub.2008.04.036.

Jones, W. & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*, 504, 427-433. doi: 10.1038/nature12715

Kelly, D.J., Quinn, P.C., Slater, A.M., Lee, K., Ge, L.Z., & Pascalis, O. (2007). The other-race effect develops during infancy – evidence of perceptual narrowing. *Psychological Science*, 18, 1084-1089. doi: 10.1111/j.1467-9280.2007.02029.x

Kontsevich, L.L., & Tyler, C.W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, 39, 2729–2737. doi:10.1016/S0042-6989(98)00285-5

Langton, S.R.H., Honeyman, H., & Tessler, E. (2004). The influence of head contour and nose angle on the perception of eye gaze direction. *Perception & Psychophysics*, 66, 752-771.

Mareschal, I., Morgan, M.J. & Solomon, J.A. (2010). Cortical distance determines whether flankers cause crowding or the tilt illusion. *Journal of Vision*, 10(8). doi: 10.1167/10.8.13.

Mareschal, I., Calder, A. J., Dadds, M. R., & Clifford, C. W. G. (2013a). Gaze categorization under uncertainty: Psychophysics and modeling. *Journal of Vision*, 13(5), 18. doi: 10.1167/13.5.18.

Mareschal, I., Calder, A. J., & Clifford, C. W. G. (2013b). Humans have an expectation that gaze is directed toward them. *Current Biology*, 23(8), 717–721. doi: 10.1016/j.cub.2013.03.030.

Mareschal, I., Otsuka, Y., & Clifford, C.W.G., (2014). A generalized tendency toward direct gaze with uncertainty. *Journal of Vision*, 14 (12). 1-9. doi: 10.1167/14.12.27.

Nelder, J.A. & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, 7, 308–313. doi: 10.1093/comjnl/7.4.308

O’Toole, A.J., Peterson, J., & Deffenbacher, K.A. (1996). An ‘other-race effect’ for categorizing faces by sex. *Perception*, 25, 669-676.

Otsuka, Y., Mareschal, I., Calder, A. J., & Clifford, C. W. G. (2014). Dual-Route Model of the Effect of Head Orientation on Perceived Gaze Direction. *Journal of Experimental Psychology: Human Perception and Performance*, in press. doi: 10.1037/a0036151

Pellicano, E., Rhodes, G., & Calder, A.J. (2013). Reduced gaze aftereffects are related to difficulties categorizing gaze direction in children with autism. *Neuropsychologia*, 51, 1504-1509. doi: 10.1016/j.neuropsychologia.2013.03.021

Rhodes, M.G. & Anastasi. J.S. (2012). The own-age bias in face recognition: a meta-analytic and theoretical review. *Psychological Bulletin*, 138, 146-174. doi: 10.1037/a0025750

Samuels, C.A. (1985). Attention to eye contact opportunity and facial motion by three-month-old infants. *Journal of Experimental Child Psychology*, 40, 105-114. doi:10.1016/0022-0965(85)90067-0

Senju, A., Yaguchi, K., Tojo, Y., & Hasegawa, T. (2003). Eye contact does not facilitate detection in children with autism. *Cognition*, 89, B43-B51. doi:10.1016/S0010-0277(03)00081-7

Senju, A., Kikuchi, Y., Hasegawa, T., Tojo, Y., & Osanai, H. (2008). Is anyone looking at me? Direct gaze detection in children with and without autism. *Brain and Cognition*, 67, 127-139. doi: 10.1016/j.bandc.2007.12.001.

Senju, A., Tucker, L., Pasco, G., Hudry, K., Elsabbagh, M., Charman, T., & Johnson, M.H. (2013). The importance of the eyes: communication skills in infants of blind parents. *Proceedings of the Royal Society*, B280, 201303436. doi: 10.1098/rspb.2013.0436

Skoczenski, A.M., & Norcia, A.M. (1998). Neural noise limitations on infant visual sensitivity. *Nature*, 391, 697-700. doi:10.1038/35630

Slepian, M.L., Weisbuch, M., Adams, R.B., & Ambady, N. (2011) Gender moderates the relationship between emotion and perceived gaze. *Emotion*, 11(6):1439-44. doi: 10.1037/a0026163.

Stoyanova, R.S., Ewbank, M.P., & Calder, A.J. (2010). "You talking to me?": Self-relevant auditory signals influence perception of gaze direction. *Psychological Science*, 21, 1765-1769. doi: 10.1177/0956797610388812

Suhrke, J., Freitag, C., Lamm, B., Fassbender, I., Poloczek, S., Teubert, M., Vohringer, I.A., Keller, H., Knopf, M., Lohaus, A., & Schwarzer, G. (2014). The other-race effect in 3-year-old German and Cameroonian children. *Frontiers in Psychology*, 5. doi: 10.3389/fpsyg.2014.00198

Symons, L.A., Hains, S.M.J. & Muir, D.W. (1998). Look at me: five-month-old infants' sensitivity to very small deviations in eye-gaze during social interactions. *Infant Behavior and Development*, 21, 531-536. doi:10.1016/S0163-6383(98)90026-1

Takahashi, K., & Watanabe, K. (2013). Gaze cueing by pareidolia faces. *i-perception*, 4(8), 490-492.

Todorovic, D. (2006). Geometrical basis of perception of gaze direction. *Vision Research*, 46, 3549-3562. doi:10.1016/j.visres.2006.04.011

Vecera, S. P., & Johnson, M. H. (1995). Eye gaze detection and the cortical processing of faces: Evidence from infants and adults. *Visual Cognition*, 2, 101–129.
doi:10.1080/13506289508401722

Vida, M.D., & Maurer, D. (2012). The development of fine-grained sensitivity to eye contact after 6 years of age. *Journal of Experimental Child Psychology*, 112, 243-256.
doi: 10.1016/j.jecp.2012.02.002

Vida, M.D., Wilson, H.R. & Maurer, D. (2014). Bandwidths for the perception of head orientation decrease during childhood. *Vision Research*, 98, 72-82.
doi:10.1016/j.visres.2014.03.005

Von Grunau, M. & Anston, C. (1995). The detection of gaze direction: A stare-in-the-crowd effect. *Perception*, 24, 1297-1313.

Webster, S., & Potter, D.D. (2008). Eye direction detection improves with development in autism. *Journal of Autism and Developmental disorders*. 38, 1184-1186. doi: 10.1007/s10803-008-0539-9