J Autism Dev Disord
DOI 10.1007/s10803-006-0232-9

ORIGINAL PAPER

Abnormal Use of Facial Information in High-Functioning Autism

**Michael L. Spezio · Ralph Adolphs ·
Robert S. E. Hurley · Joseph Piven**

Abstract Altered visual exploration of faces likely contributes to social cognition deficits seen in autism. To investigate the relationship between face gaze and social cognition in autism, we measured both face gaze and how facial regions were actually used during emotion judgments from faces. Compared to IQ-matched healthy controls, nine high-functioning adults with autism failed to make use of information from the eye region of faces, instead relying primarily on information from the mouth. Face gaze accounted for the increased reliance on the mouth, and partially accounted for the deficit in using information from the eyes. These findings provide a novel quantitative assessment of how people with autism utilize information in faces when making social judgments.

Keywords Social cognition · Emotion · Eyetracking · Bubbles · Facial information

M. L. Spezio · R. Adolphs (✉)
Division of Humanities and Social Sciences, 228-77,
California Institute of Technology, Caltech, Pasadena, CA
91125, USA
e-mail: radolphs@hss.caltech.edu

M. L. Spezio · R. Adolphs
Computation and Neural Systems, California Institute of
Technology, Pasadena, CA 91125, USA

R. S. E. Hurley · J. Piven
Neurodevelopmental Disorders Research Center,
University of North Carolina, Chapel Hill, NC 27599, USA

M. L. Spezio
e-mail: mlspezio@hss.caltech.edu

Introduction

Autism is a neurodevelopmental disorder strongly characterized by deficits in social interaction and impaired understanding of the mental states of others (Baron-Cohen, 1997; Frith & Frith, 1999; Kanner, 1943; Siegel, Vukicevic, & Spitzer, 1990), a dysfunction that persists even in people with autism who have IQs in the normal range. Because high-functioning children and adults with autism show (Baron-Cohen et al., 1999; Buitelaar, van Engeland, de Kogel, de Vries, & van Hooff, 1991; Carpenter, Pennington, & Rogers, 2002; Castelli, Frith, Happe, & Frith, 2002; Loveland, Pearson, Tunali-Kotoski, Ortegon, & Gibbs, 2001; Ozonoff & Miller, 1995; Rogers, 2000; Rogers, Hepburn, Stackhouse, & Wehner, 2003) and report (Gilpin, 2002; Grandin, 1996) difficulties in social judgment (e.g., understanding others' emotions, deciding on appropriate social behaviors, etc.), a main focus of autism research has been to understand how people with autism process salient social cues, notably from faces. There has been a considerable amount of work using static faces (i.e., photographs) to investigate social judgments (Capps, Yirmiya, & Sigman, 1992; Celani, Battacchi, & Arcidiacono, 1999; Critchley et al., 2000; van der Geest, Kemner, Verbaten, & van Engeland, 2002a; Grelotti, Gauthier, & Schultz, 2002; Joseph & Tanaka, 2003; Langdell, 1978; Ogai et al., 2003; Trepagnier, Sebrechts, & Peterson, 2002; Volkmar, Sparrow, Rende, & Cohen, 1989; Weeks & Hobson, 1987) and gaze fixation behavior (Buitelaar et al., 1991; Carpenter et al., 2002; van der Geest et al. 2002a; van der Geest, Kemner, Camfferman, Verbaten, & van Engeland, 2002b; Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Pedersen, Livoir-Petersen, & Schelde, 1989; Pelphrey et al., 2002; Trepagnier et al., 2002; Volkmar et al., 1989). A seminal study by Langdell (1978) showed that children with autism were better than controls at judging facial identity based on partially presented features of the face such as the eye or mouth regions. Younger children showed particularly heavy reliance on the mouth region. Recently, Joseph and Tanaka (2003) demonstrated that high-functioning children with autism were much better at judging facial identity from the mouth alone than from the eyes alone, and in comparison to age- and IQ-matched controls, were impaired at judging facial identity from the eyes alone. Studies of face gaze in autism have generally been consistent with these abnormalities in facial information processing. High-functioning children and adults with autism have been found to allocate more gaze to the mouth than to the eyes during viewing of dynamic and static facial stimuli (Klin et al., 2002; Pelphrey et al., 2002).

A critical open question is the degree to which abnormal face gaze might contribute to impairments in the use of facial information during social judgment, as has been demonstrated in the case of a neurological patient with amygdala damage (Adolphs et al., 2005). The difficulty in establishing this link is that gaze is only a first stage in a series of visual processing steps that eventually culminates in the social judgments measured. These additional processing stages can influence information use and performance independently of precise direction or amount of gaze. For instance, studies of overt visual attention have shown that visual processing can vary independently of any change in fixation (Triesch, Ballard, Hayhoe, & Sullivan, 2003; Turatto, Angrilli, Mazza, Umilta, & Driver, 2002).

Here we probed how people are able to use information from regions of the face in order to judge emotion, and how they simultaneously fixated regions of the face. Our approach, utilizing the ‘‘Bubbles’’ method (Gosselin & Schyns, 2001), combines the use of static facial stimuli with a measure of the facial information that people with autism actually use in social judgment. This approach was used in a previous study to determine that focal lesions of the amygdala specifically affect how information from the eyes is processed in emotion judgments of fear from faces (Adolphs et al., 2005).

The ‘‘Bubbles’’ method yields those regions of a face that are strongly associated with making a specified judgment about the face. During ‘‘Bubbles,’’ a given trial shows only randomly revealed areas of the face, determined by the number of ‘‘bubbles, or Gaussian holes in a mask covering the underlying, or base, image. The more bubbles there are, the more area of the face is revealed to a viewer. The viewer then makes a judgment based on what is revealed. Averaging performance across all the trials yields an image, called the ‘‘diagnostic image,’’ that depicts which areas of the face, on average, contributed most to making the judgments. For example, if we asked for judgments of ear size using static facial images, the analysis of a ‘‘Bubbles’’ experiment would yield an image prominently showing the ears but missing the eyes and mouth. So what is seen in a ‘‘Bubbles’’ diagnostic image is the information viewers rely on to make judgments about the face. We combined eyetracking with ‘‘Bubbles’’ in order to move beyond making inferences about facial information processing from gaze alone.

Two primary interpretations relating face gaze and social cognition could account for differences in information use we might observe using this procedure. If participants with autism were to differ from controls both in their use of information from certain facial regions as well as in their gaze to those same regions, then we could conclude that face gaze strongly contributes to the differences in use of facial information. However, if face gaze to a region were no different from that of controls while the use of facial information were to differ, then face gaze would not be a sufficient explanation for the observed differences in facial information processing of that region. In this case, we would need to conclude that the participants with autism process faces abnormally, above and beyond their fixations to them.

Methods

Research Participants

All research methods were conducted with the approval of either the Institutional Review Board at the California Institute of Technology or the Institutional Review Board at the University of North Carolina. Nine high-functioning male participants with autism (HFA) were recruited through the Subject Registry of the Neurodevelopmental Disorders Research Center at the University of North Carolina, where they were tested. All HFA participants met DSM-IV/ICD-10 diagnostic criteria for autism, and all met the cutoff scores for autism on both the Autism Diagnostic Interview (LeCouteur, Rutter, & Lord, 1989) and the Autism Diagnostic Observation Schedule (Lord et al., 1989). We assessed IQ for all participants using the Wechsler Abbreviated Scale of Intelligence (WASI$^{\text{TM}}$). The HFA group had a mean age of 23 years (20, 22, 21, 26, 20, 20, 18, 40, 20), and mean IQ values of 109 verbal (108, 77, 122, 74, 120, 130, 87, 131, 134), 104 performance (111, 118, 104, 97, 91, 119, 82, 94, 125), and 107 full scale (111, 96, 115, 83, 106, 128, 83, 112, 133). Ten male participants were enrolled as controls (C) and tested at Caltech with the same protocols as were used for the HFA participants. Control participants had no history of neurological or psychiatric disease or pervasive developmental disorder or other evidence of developmental disability, or family history of autism. Controls had a mean age of 28 years (20, 20, 22, 22, 22, 40, 39, 34, 32, 35), and mean IQ values of 101 verbal (83, 76, 81, 123, 104, 109, 121, 105, 95, 117), 111 performance (93, 106, 98, 119, 118, 106, 119, 109, 121, 119), and 106 full scale (86, 88, 88, 125, 111, 109, 124, 108, 108, 118). There was no significant difference between the HFA group and controls in age, or in verbal, performance, or full-scale IQ ($P > 0.1$ for each comparison, t-tests on independent samples). All participants had normal or corrected to normal vision at testing time. Concurrent eyetracking measurements during the Bubbles task (see below) were performed only for the first eight HFA participants, due to logistical and time constraints. There were no statistically significant differences in age or IQ between this subset of the HFA group and the full set of controls. Concurrent eyetracking measurements during emotion judgment from unfiltered faces (see below) were performed only for the first eight HFA participants and for only five of the ten control participants. There were no statistically significant differences in age or IQ between these participant subgroups.

Procedures

All eyetracking data and button responses were recorded using the Eyelink II head-mounted eyetracking system (SR Research, Hamilton, Ontario, Canada). Eyetracking data were recorded at either 250 or 500 Hz. New nine-point calibrations and validations were performed prior to the start of each experiment in a participant's session. Accuracy in the validations typically was better than 0.5° of visual angle. Experiments were run under WindowsXP (Microsoft Inc.) in Matlab (Mathworks Inc., Natick, MA, USA) using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and the Eyelink Toolbox (Cornelissen, Peters, & Palmer, 2002).

Emotion judgment of facial expressions in the "Bubbles" task used faces with randomly revealed regions as previously described (Gosselin & Schyns, 2001). Briefly, on each trial, a randomly selected base facial image was first decomposed into a six-level Laplacian pyramid using the Simoncelli steerable pyramid toolbox for Matlab (Portilla & Simoncelli, 2000) with a Gaussian filter subtending 1° of visual angle (11 w × 11 h). Levels one through five were then filtered with a number of bubbles whose centers were randomly distributed across the image. After filtering, levels one through five were combined with a standard background corresponding to the sixth level, and the resulting image was presented. The number of bubbles was adjusted for each participant on a trial-by-trial basis in order to maintain performance accuracy of response near 80% correct. Note that bubbles were allowed to overlap, increasing the amount of the face revealed beyond the size of a single bubble. Base stimuli (256 w × 256 h; pixel units) were cropped from four normalized Ekman faces (Ekman & Friesen, 1976), each of a different posing participant, and balanced for gender and facial expression (two fearful, two happy). Images were normalized for magnitude across all spatial frequencies and centrally displayed using a monitor resolution of 640 w × 480 h (pixel units) on a 15.9 in. w × 11.9 in. h monitor, at an eye-to-screen distance of approximately 31 in., thus subtending 11.3° of horizontal visual angle.

A given trial lasted the time it took participants to decide whether the face showed fear or happiness (Adolphs et al., 2005), for a maximal decision time of 10 s following image onset. Participants were asked to judge whether the bubbled face they saw was afraid or happy, by pushing a button. All participants completed 512 trials. On every fifth trial, a circular annulus was centrally displayed and participants were given an opportunity to rest. When they decided to continue, they fixated the annulus and simultaneously pressed a key. This advanced the experiment to the next trial and allowed the system to correct for any drift in eyetracking accuracy. Participants were instructed to decide as quickly as possible and to always make a decision, even if it was a best guess.

Emotion judgment of unfiltered faces used 46 standard Ekman faces balanced for gender and identity (6 different identities for each basic emotion of happiness, sadness, fear, anger, surprise, and disgust, and 10 different identities for neutral). Images (512 w × 768 h) were normalized for overall intensity and centrally displayed using a monitor resolution of 1,280 w × 1,024 h (pixel units) on a 15.9 in. w × 11.9 in. h monitor, at an eye-to-screen distance of approximately 31 in., thus subtending ~11° of horizontal visual angle. Images were displayed for 1 s followed by a list of basic emotion words. Participants were asked to name the emotion seen in the face.

‘‘Bubbles’’ Analysis

‘‘Bubbles’’ data were analyzed as previously described (Gosselin & Schyns 2001; Schyns, Bonnar, & Gosselin, 2002), with some modification. Analyses determined which regions of the face associated with correct emotion judgments. To draw an association between facial regions and correct emotion judgments, we summed the trial-specific five-level bubbles masks across all correct trials and across all incorrect trials, yielding a ‘‘correct’’ and an ‘‘incorrect’’ bubbles mask for each level. We then subtracted, for each spatial frequency level, the normalized incorrect from the normalized correct mask, resulting in a difference mask. In order to select regions of statistically significant difference for the difference mask, we converted all pixel values into *Z*-scores relative to that mean and standard deviation. The statistical analyses of the *Z*-scored classification image proceeded by a recently developed method (Chauvin et al., submitted; http://mapageweb.umontreal.ca/gosselif/Stat4Ci_rev.pdf) that uses the same approach as that used for the statistical analysis of significant clusters of activation in fMRI and PET data (Friston, Worsley, Frackowiak, Mazziotta, & Evans, 1994). After smoothing with a Gaussian filter having sigma = 5, we subjected this *Z*-scored classification image to cluster tests, setting a threshold $t = 2.5$ and a significance $P = 0.001$. This resulted in a diagnostic image, showing which features of the face were relied upon most during the behavioral task.

Further Quantification of Individual Participants’ Reliance on the Eyes and Mouth

In addition to the group analyses described above, we sought to quantify each individual’s reliance on the eyes and mouth during emotion judgment. We estimated the effective strength of the appearance of each region of interest (ROI, i.e., eyes and mouth) in an individual’s diagnostic image, using a metric knows as the Structural SIMilarity (SSIM) index. SSIM was developed by Wang, Bovik, Sheikh, and Simoncelli (2004) as a quantitative estimate of the similarity between two images that corresponds closely to similarity judgments by human observers. SSIM values were calculated between each individual’s diagnostic image and the corresponding base image, for each specified region. We also calculated an eye-to-mouth quality ratio (EMQR) by calculating the SSIM for each eye, taking the maximum of both values and dividing this by the SSIM for the mouth. This yielded a quantifiable estimate of each individual’s relative reliance on the eyes and mouth in making emotion judgments. The greater a participant’s EMQR, the greater the participant’s relative reliance upon information in the eye region(s) compared to the mouth region.

Analysis of Performance and Gaze Behavior

Eyetracking data were analyzed for fixations using the Eyelink DataViewer (SR Research, Hamilton, Ontario, Canada). In discriminating fixations, we set saccade velocity, acceleration, and motion thresholds to 30°/s, $9{,}500°/s^2$, and 0.15°, respectively. Measures of face gaze included fixation number (i.e., the total number of fixations within an area, independent of previous fixation area) and fractional dwell time (i.e., the time during a given trial spent fixating a given area divided by the total time between image onset and response).

Regions of interest were drawn for each facial image, using the drawing functions within the DataViewer. We used nine ROIs in all: right eye (around the white of the right eye), left eye (around the white of the left eye), right eye region (including the right eye and the eye socket around it), left eye region (including the left eye and the eye socket around it), nose, mouth, face, right eyebrow, and left eyebrow. The designations right and left are anatomical, and not from the perspective of the viewer. See Fig. 3g for a visual depiction of regions.

For Bubbles trials, face gaze was analyzed by including all fixations on correct trials that began within a time window between 50 ms following stimulus onset and the response. For unfiltered Ekman faces, face gaze was analyzed by including all fixations that began within a time window between 50 ms following stimulus onset and stimulus offset at 1,000 ms.

Results

Emotion Judgment with ‘‘Bubbles’’ Faces

To determine those areas of the face that were in fact used by participants in order to judge the emotion shown in the stimuli, we showed random areas of the face, using the ‘‘Bubbles’’ technique, and analyzed the areas of the face that were revealed as a function of participants’ accuracy. We used four different Ekman faces, two female and two male, counterbalanced for showing happiness or fear. We chose happy and fearful faces based upon an earlier study involving focal amygdala lesions (Adolphs et al., 2005), and because the range of the difficulty in judging basic emotions from the face is captured by happy and fearful faces.

The number of bubbles per facial image was adjusted on a trial-by-trial basis in order to maintain an accuracy of response in the range of 80%, and in fact HFA participants ($n = 9$) and controls ($n = 10$) showed the same performance accuracy (HFA: 82 ± 3%; Con: 80 ± 5%; M ± SD). There was also no difference in mean reaction time (HFA: 1.53 ± 0.64 s; Con: 1.46 ± 0.53 s), maximal reaction time (HFA: 7.44 ± 2.53 s; Con: 7.52 ± 2.25 s), median reaction time (HFA: 1.16 ± 0.34 s; Con: 1.04 ± 0.29 s), or the number of bubbles per facial image (HFA: 62 ± 29; Con: 52 ± 22; $P > 0.1$ for each comparison, Wilcoxon rank-sum tests).

We assessed which regions of the face were used to judge emotion from the ‘‘Bubbles’’ stimuli, across all trials. The diagnostic images shown in Fig. 1a, b show that while the participants in the HFA group used information from both the eyes and the mouth, they did so differently than controls. Difference images (Fig. 1c, d) show that HFA participants made significantly less use of information from the eyes and more use of information from the mouth. Note that these difference images visually display regions of the face whose use between groups differed at statistically significant levels (see Sect. ’’Methods’’ for details).

To quantify each individual’s relative reliance on the eyes and mouth during emotion judgment, we calculated EMQRs from each participant’s diagnostic image (see Sect. ’’Methods’’). The higher the EMQR, the greater was the reliance on the eyes, relative to the mouth, in performing the emotion judgment task. There was a group difference in EMQR (HFA: 0.61 ± 0.38; Con: 1.18 ± 0.55; $P < 0.03$, Wilcoxon rank-sum test), and six of nine HFA participants (67%) had EMQRs that were lower than one standard deviation below the mean for controls. Only one control participant showed an EMQR this low (Fig. 2a). Thus, patterns of facial feature usage within and between groups that are shown in Fig. 1 are borne out by this further quantification of the data from each individual. Note that the difference in individual use of the eyes was the main contributor to differences in the EMQR, as seen in Fig. 2b, c. There was a group difference in SSIM for the eye that was maximally used (HFA: 0.25 ± 0.18; Con: 0.43 ± 0.15; $P < 0.05$) but not for the mouth ($P > 0.1$; Wilcoxon rank-sum tests).

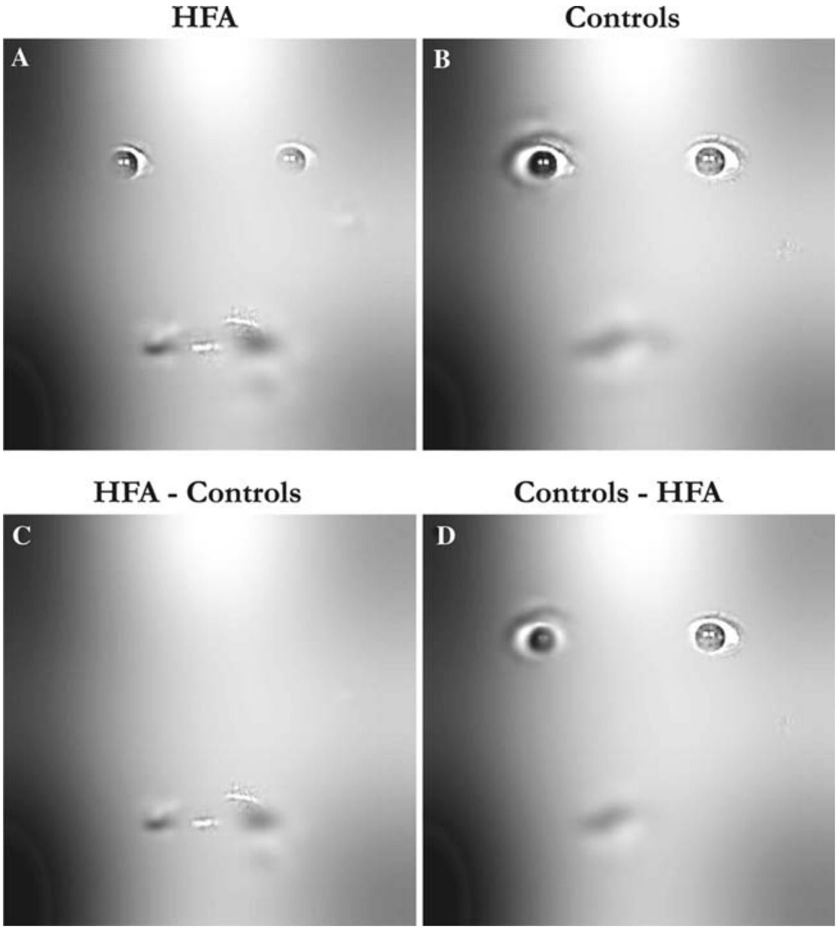

Fig. 1 Use of facial information when judging emotion from ‘‘Bubbles’’ faces. Shown at the *top* is the facial information (i.e., the *key facial regions*) used by the high-functioning participants with autism (HFA group) (**a**) and controls (**b**) to judge emotion. Subtracting these images from each other reveals the facial information that was used more by the HFA group than by controls (**c**) and more by the controls than by the HFA group (**d**). Note that these images depict statistically thresholded differences; the facial features shown are thus those that differed significantly ($p < 0.001$ with a cluster threshold $t = 2.5$) in their use between the two subject groups

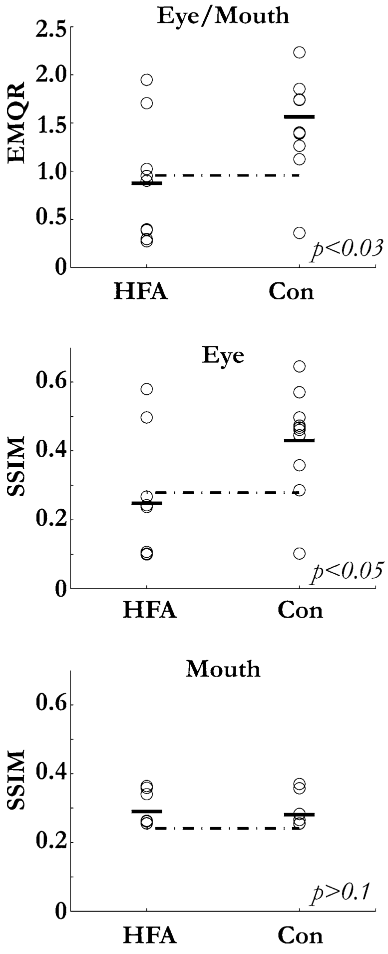

Fig. 2 Individuals' use of the eyes and mouth when judging emotion from informationally constrained faces. To determine how well the group patterns in Fig. 1 characterized each group, we calculated each individual's use of the eye and mouth regions using structural similarity (SSIM) comparisons between the individual's diagnostic image and the corresponding base image (see Sect. "Methods"). For each plot (**a–c**), each *circle* is one individual participant, the *dark horizontal bars* are the mean values for each group, and the *dotted horizontal lines* show one standard deviation below the higher mean. Shown also are individual *p* values resulting from Wilcoxon rank-sum tests of differences in group means. A. eye-to-mouth quality ratio (EMQR). EMQR values quantify an individual's relative use of the eyes compared to the mouth in emotion judgments. Six of nine HFA participants showed EMQR values that were lower than one standard deviation below the mean for controls, while this was seen for only one of the control participants. **b** SSIM for the eye that was maximally used during emotion judgment. **c** SSIM for the mouth

Face Gaze with "Bubbles" Faces

The most straightforward hypothesis to explain these differences in the use of facial features during emotion judgments (Fig. 1, 2) is that HFA participants actually looked more at the mouth and less at the eyes. We tested this hypothesis by performing eyetracking measurements in eight of the nine HFA participants during the emotion judgment task and then comparing their facial fixation behavior to that of controls. After confirming that this subset of eight HFA participants showed the same pattern of facial information use as the full group, we calculated the mean fractional dwell time (Fig. 3a–c).

We focused on the three facial regions that showed significant differences in the diagnostic images: the right eye, the left eye, and the mouth (Fig. 1). There was a significant difference in gaze to the mouth (Fig. 3c; $P < 0.005$; Wilcoxon rank-sum tests), such that the HFA group spent a greater proportion of a trial looking at (HFA: 0.176 ± 0.086; Con: 0.049 ± 0.034; $P < 0.005$) and made more fixations to (HFA: 299 ± 109; Con: 111 ± 68; $P < 0.005$) the mouth, compared to controls. Additionally, the analysis of fixations to the right eye also reveals that the HFA group spent less time fixating the right eye (HFA: 0.011 ± 0.011; Con: 0.042 ± 0.032; $P < 0.05$), when the top outlier in the HFA group was removed (Fig. 3a, see figure legend). However, there was no group difference in the time spent fixating the left eye (HFA: 0.035 ± 0.023; Con: 0.048 ± 0.030; $P > 0.1$).

Note that these group differences in face gaze behavior are not due to overall differences in the revealed features that each subject saw during the "Bubbles" task. As shown in Fig. 3d–f, both groups (HFA, solid line; controls, dotted line) received the same number of bubbles at each spatial frequency in the regions of the right eye (Fig. 3d), left eye (Fig. 3e), and mouth (Fig. 3f). The same was true for all other regions (not shown). Moreover, each individual received, over the course of a "Bubbles" task, essentially the same coverage of bubbles in each feature (Supplementary Figs. 1, 2), as would be expected given that the "Bubbles" method uses a homogenous random sampling of the image. While each individual trial on the task samples the face differently, this sampling is entirely random; hence it is unrelated to the group of participants (HFA or control), and accumulates to a nearly identical cumulative sampling of the face over the large number of trials we used. Thus, the observed differences in face gaze (Fig. 3a–c), calculated over an entire "Bubbles" session, cannot be the result of systematic stimulus differences between groups.

These findings support the hypothesis that the differences observed in using facial information during emotion judgment were due in large part to differences in face gaze. Differences in the use of the mouth can be attributed entirely to differences in face gaze. The HFA group spent significantly more time fixating the mouth than did controls (Fig. 3c), and used the mouth significantly more than controls during emotion judgment (Fig. 1c). For the

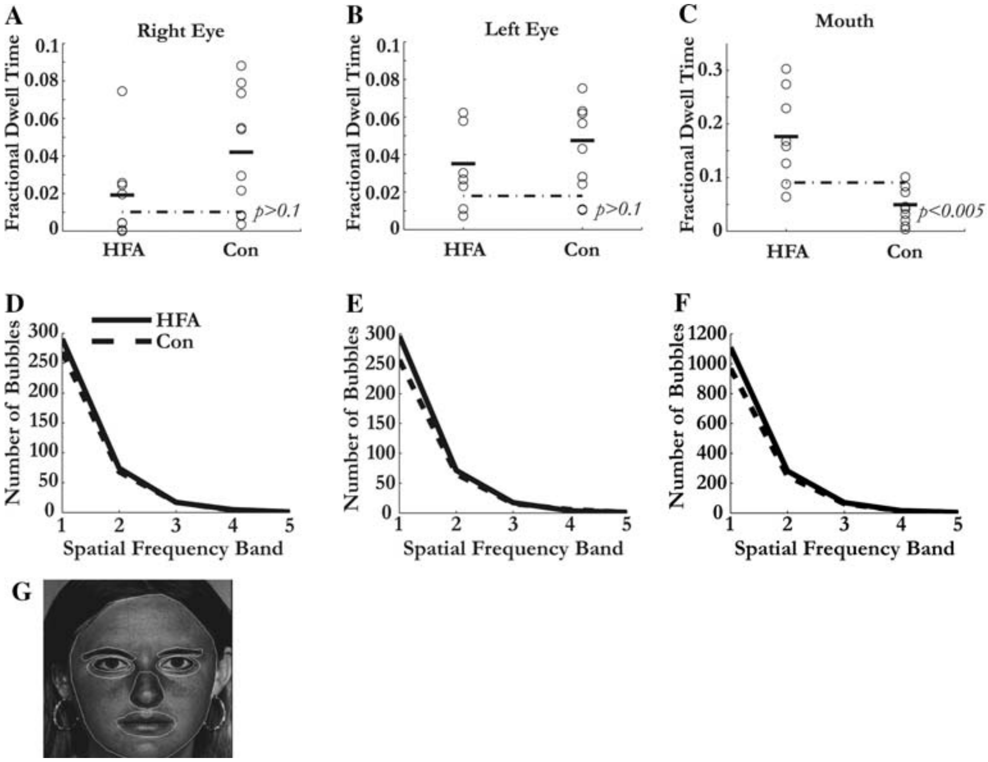

Fig. 3 Differences in use of facial information are not fully explained by differences in face gaze. To determine if the differences in the use of facial information (Fig. 1) could be explained by differences in face gaze, we simultaneously performed eyetracking measurements during the Bubbles task. We compared a subset (eight of nine) of high-functioning participants with autism (HFA) to the control group. The subset of HFA participants showed the same pattern of differences in information use as did the full group. *Panels a–c* show mean fractional dwell time for the eyes and mouth, across all trials. Fractional dwell time is calculated for each trial by dividing the total time of fixation within a given region by the total time between image onset and response. Each *circle* is one individual participant, the *dark horizontal bars* are the mean values for each group, and the *dotted horizontal lines* show one standard deviation below the higher mean. Shown also are individual *p* values resulting from Wilcoxon rank-sum tests of differences in means for the HFA and control (Con) groups. There was a significant difference in gaze to the mouth (**c**), and removing the HFA participant with the highest fractional dwell time in the eye regions resulted in a significant group difference for the right eye [(**a**), $p < 0.05$]. However, there was no group difference in fixation to the left eye [(**b**), $p > 0.1$]. These findings for face gaze are not due to stimulus differences, as shown in (**d**)–(**f**). Each *panel* depicts the group mean of the number of bubbles (*ordinate*) at a given spatial frequency (*abscissa*) in a given facial region. Data for the HFA and control groups are shown with *solid* and *dashed* lines, respectively. An example of facial region definition is shown in (**g**) (see Sect. ''Methods''), overlaid on a standard Ekman face

right eye, the results are also consistent with the notion that gaze accounted for the difference in use (Figs. 1d, 3a). However, HFA participants were no different than controls in the time spent fixating the left eye, yet they showed significantly less use of both eyes (Figs. 1d, 2b). Thus, differences in fixation account for most of the differences in actual use of information from facial regions during emotion judgment, but do not fully account for all abnormal information use in the case of the eyes.

Emotion Judgment and Face Gaze with Unfiltered Whole Faces

In order to obtain a standard to which to compare our findings for the ''Bubbles'' faces, we turned to an emotion judgment task using unfiltered Ekman faces. We asked participants (five controls, eight HFA) to judge the emotion in standard Ekman faces showing either one of six basic emotions (happiness, sadness, fear, anger, surprise, disgust) or a neutral expression (see Sect. ''Methods''). We simultaneously measured their response accuracy and their face gaze. Given that the HFA group performed normally on the more difficult ''Bubbles'' task, that they were quite well-matched on IQ to the controls, that all of our HFA participants had performed simple emotion judgments from faces in past studies, and that all had undergone intensive social gaze training using static faces, we did not expect to see any group differences in either accuracy of emotion judgment or in face gaze. In fact, we found no group differences in accuracy (i.e., percent correct) to any of the basic emotions ($P > 0.1$ for all comparisons; *t*-test for independent samples). Nor did we find group differences in fractional dwell time to the right eye (HFA: 0.020 ± 0.013; Con: 0.020 ± 0.025); left eye (HFA: 0.015 ± 0.028; Con: 0.021 ± 0.037); right eye region (HFA: 0.064 ± 0.035; Con: 0.046 ± 0.057); left eye region (HFA: 0.036 ± 0.052; Con: 0.042 ± 0.066); nose (HFA: 0.113 ± 0.050; Con: 0.206 ± 0.096); mouth (HFA: 0.062 ± 0.035; Con: 0.054 ± 0.029); right eyebrow (HFA: 0.008 ± 0.008; Con: 0.002 ± 0.003); or left eyebrow (HFA: 0.001 ± 0.002; Con: 0.001 ± 0.002) ($P > 0.1$ for all comparisons; Wilcoxon rank-sum test).

A comparison of the face gaze data obtained from the different stimuli shows that, in general, the HFA group increased gaze to the mouth when shown ‘‘Bubbles’’ stimuli compared to whole face stimuli, whereas controls increased gaze to the eyes when shown ‘‘Bubbles’’ stimuli compared to whole face stimuli. As noted earlier, these group differences are not due to any group difference in revealed features during ‘‘Bubbles’’ (see Fig. 3d–f). We believe they are due to the increased difficulty in judging the ‘‘Bubbles’’ faces, compared to whole faces, eliciting different compensatory behaviors in each group, an issue we take up in more detail in Sect. ’’Discussion.’’

Discussion

Bubbles, Face Gaze, and the Use of Facial Information

This is the first report to assess directly how information from different features of the face is used by people with autism during judgment of basic emotions. We isolated a specific face processing impairment in people with autism by employing a novel approach to facial information processing, simultaneously controlling for effects of IQ, performance accuracy, and reaction time. We showed that individuals with autism were strongly distinguished from controls in the features they relied upon most while making emotion judgments. The HFA group showed a strikingly decreased use of information from the eye regions and a marked reliance upon information from the region of the mouth, compared to controls.

The different strategy in how the HFA group used information from the eyes and mouth is especially striking given that there were no overall performance differences between groups: the HFA group performed at the same level of accuracy, and with the same reaction time, as did controls on the task. Accuracy on the ‘‘Bubbles’’ task was, of course, determined by the stimulus display software to be near 80% correct, yet one might have expected differences in the number of bubbles or in the reaction time necessary to achieve this level of performance. The fact that we observed no emotion judgment differences in the ‘‘Bubbles’’ task suggests that the HFA group productively used the information present in the mouth region. That is, they achieved the same performance level as did controls, but through using a different face processing strategy. The emotional expressions used in this experiment—fear and happiness—do in fact differ in the mouth region. Based on our results here, it is likely that the ‘‘Bubbles’’ method would reveal a performance deficit in HFA participants in judgments of emotional expressions that differed clearly only in the eye regions and not in the area of the mouth, or in emotion judgments from stimuli that only show the eyes [e.g., the task of Baron-Cohen, Wheelwright, Hill, Raste, & Plumb (2001)].

One caveat in interpreting these findings is the possibility that the ‘‘Bubbles’’ method, which reveals only certain areas of an object on any given trial, alters strategies of visual processing. This possibility has indeed been raised in the literature and demonstrated for a case of simple object processing (Murray & Gold, 2004). However, it has also been demonstrated that the ‘‘Bubbles’’ method does not elicit an altered visual processing strategy for faces in emotion judgment tasks (Gosselin & Schyns, 2004). Thus we consider it likely that our findings reflect facial information processing strategies typically employed by the HFA participants and controls.

We also consider it unlikely that our face gaze findings for the ‘‘Bubbles’’ faces are artifacts of the filtered stimuli. We have shown that there are no group differences in how much the eyes and mouth were revealed across ‘‘Bubbles’’ sessions (Fig. 3d–f), and we have also determined that there are no individual differences in this regard, (Supplementary Figs. 1, 2). Comparing the fractional dwell time in the eyes and mouth for unfiltered and ‘‘Bubbles’’ faces shows that controls increased their gaze to the eyes and HFA participants increased their gaze to the mouth during emotion judgment with ‘‘Bubbles’’ faces. It is highly likely that this was due to the increased difficulty of the Bubbles task over emotion judgment with unfiltered faces.

Unfiltered Whole Faces

Several aspects of the data for whole faces are noteworthy. First, we found no impairment in performance accuracy in judging the emotion from whole faces in the HFA group. This was not entirely surprising, given that the HFA group also performed at a normal level of accuracy on the much more difficult ''Bubbles'' task. Nonetheless, it is at odds with some reports in the literature (Capps et al., 1992; Celani et al., 1999), although it is consistent with others (Ogai et al., 2003). While we did not study the underlying factors that may have contributed to the essentially normal performances of the HFA group on emotion judgment accuracy, several factors may have contributed. The HFA group was very well-matched in IQ and age to our control group, participants in the HFA group previously had participated in other studies requiring emotion judgment with faces, and all HFA participants had received extensive, long-term, training in social gaze and emotion judgment as part of their interventional therapy. It will be important to accumulate future studies that can begin to map variance in the emotion recognition performances of people with autism onto variances in underlying factors such as intervention type, IQ, or other aspects of subject heterogeneity.

Perhaps more puzzling is our finding that the HFA group also did not show any abnormality in fixations onto the whole faces, unlike their abnormal fixations onto the ''Bubbles'' stimuli. One possible explanation might simply be that this task lacked the statistical power to detect such abnormalities: whereas the ''bubbles'' task measured fixations for each participant on a total of 512 trials, the whole faces task comprised only 46 trials. Perhaps a much larger number of trials would reveal subtle differences in fixations also in the whole faces. Nonetheless, it seems clear that any abnormalities in fixations for the whole faces would be much smaller in magnitude than for the ''Bubbles'' faces. Our explanation of this difference is that the ''Bubbles'' task is much more difficult, and the sparse stimuli that it uses accentuate an ability to fixate onto relevant features in order to extract the maximal information from the few features that are revealed. In our view, this further justifies the use of ''Bubbles'' method as a sensitive probe into visual information processing that can reveal abnormalities in social information processing that may not be apparent with richer stimuli or on easier tasks.

Neural Substrates

We noted that the most obvious explanation for our finding that the HFA group relied less upon the eyes and more upon the mouth than did controls would be that the HFA group simply fixated the eyes less and the mouth more. Our findings support this hypothesis to a degree, consistent with recent neuroimaging work in people with autism (Dalton et al., 2005). Yet differences in face gaze did not fully account for the difference in use of information from the left eye, suggesting that there may be additional processing stages that are abnormal in autism. We suggest that the brains of people with autism treat facial information differently, even when the visual input is the same. This hypothesis could be tested directly in future studies by adding functional imaging to the experimental approach described here. Experiments in this vein would likely better distinguish the precise roles of brain regions previously implicated in autism, such as the fusiform gyrus, the superior temporal cortex, and the amygdala (Grelotti et al., 2002; Hadjikhani et al., 2004; Pelphrey, Adolphs, & Morris, 2004).

The fusiform face area is known to be hypoactive in response to facial stimuli in autism, compared to responses in matched controls (Schultz et al., 2003). Dalton et al. (2005) proposed that the cause of this observed hypoactivation is a failure in autism to make direct eye contact. Indeed, they showed a positive correlation between the duration of direct eye contact and percent signal change in the right anterior fusiform gyrus that was present only for the autistic group and not for controls. Interestingly, in comparing the autistic and control groups in that study, the percent signal change *difference* between the autistic and control groups was partially independent of eye gaze. Participants with autism who showed average eye fixation durations similar to those shown by controls nevertheless showed a percent signal change reduced by nearly one half [Dalton et al. (2005), Fig. 7]. Taken together, our findings are consistent with those of Dalton et al. (2005) and with our proposal that the facial information processing in brains of people with autism is abnormal, even when direct eye gaze is the same.

Additional insight into the possible role of the amygdala in autism is gained by comparing our results to those of a previous study with a patient, S.M., who has bilateral focal amygdala lesions (Adolphs et al., 2005). The study with S.M. used the same stimuli and task as used here, though without concurrent eye-tracking during the ''Bubbles'' task. A comparison reveals an interesting similarity between S.M. and our HFA group, in that both showed significantly decreased use of the eyes in comparison to controls. While this might suggest that the abnormal use of eye information seen for the HFA group was due to amygdala dysfunction, two important differences between our results and those in Adolphs et al. (2005)

warrant caution in drawing this conclusion. First, the HFA group relied more on both eyes than did S.M. While S.M.'s diagnostic image showed almost no use of information from the eyes, the HFA group did show use of the eyes, although that use was greatly reduced. Second, the HFA group was much more strongly dependent than S.M. on information from the mouth. Yet there is an area of overlap between our study and that involving S.M. When S.M. foveated the eyes in a face after being expressly directed to do so, S.M.'s ability to judge fear from whole faces was restored. Our results suggest that directed eye-to-eye gaze, if it overcame gaze to the mouth, might partially restore the use of information from the eyes in adults with high-functioning autism. However, given our findings, it is not at all clear that correcting face gaze in autism would fully overcome deficits in the use of facial information. In summary, the comparison with S.M. suggests that while amygdala dysfunction may play a role in the abnormal strategy for using facial information we describe here, impairments in wider neural networks for social cognition that include structures in addition to the amygdala would be required to account fully for the deficits we report in the HFA participants (Baron-Cohen et al., 1999, 2000; Frith, 2003; Pelphrey et al., 2004; Schultz et al., 2003).

Our findings have implications for autism intervention programs that focus on normalizing face gaze. Face gaze normalization often has two stated goals, one of which is to help persons with autism display social behavior that is more conducive to engaging others (e.g., eye-to-eye contact), and another which is to help persons with autism make use of facial cues necessary for social fluency. The results here are consistent with this second goal in large part, but they do suggest that obtaining improved use of facial information from face gaze normalization may be more complicated to achieve than originally thought.

Acknowledgments The authors would like to thank the participants and their families for making this study possible, Dr Frédéric Gosselin for helpful advice in using the "Bubbles" method, and Dr Fulvia Castelli and anonymous reviewers for helpful comments. We gratefully acknowledge the support of NIMH STAART Center funding and a grant from the Cure Autism Now Foundation.

References

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433*(7021), 68–72.

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. Cambridge: MIT.

Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C., & Williams, S. C. (2000). The amygdala theory of autism. *Neuroscience Biobehavioral Reviews, 24*(3), 355–364.

Baron-Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A., et al. (1999). Social intelligence in the normal and autistic brain: An fMRI study. *The European Journal of Neuroscience, 11*(6), 1891–1898.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*(2), 241–251.

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

Buitelaar, J. K., van Engeland, H., de Kogel, K. H., de Vries, H., & van Hooff, J. A. (1991). Differences in the structure of social behaviour of autistic children and non-autistic retarded controls. *Journal of Child Psycholology and Psychiatry, 32*(6), 995–1015.

Capps, L., Yirmiya, N., & Sigman, M. (1992). Understanding of simple and complex emotions in non-retarded children with autism. *Journal of Child Psychology and Psychiatry, 33*(7), 1169–1182.

Carpenter, M., Pennington, B. F., & Rogers, S. J. (2002). Interrelations among social-cognitive skills in young children with autism. *Journal of Autism and Developmental Disorders, 32*(2), 91–106.

Castelli, F., Frith, C., Happe, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain, 125*(Pt 8), 1839–1849.

Celani, G., Battacchi, M. W., & Arcidiacono, L. (1999). The understanding of the emotional meaning of facial expressions in people with autism. *Journal of Autism and Developmental Disorders, 29*(1), 57–66.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The eyelink toolbox: Eye tracking within MATLAB and the psychophysics toolbox. *Behavioral Research Methods, Instrumentation and Computers, 34*, 613–617.

Critchley, H. D., Daly, E. M., Bullmore, E. T., Williams, S. C., Van Amelsvoort, T., Robertson, D. M., et al. (2000). The functional neuroanatomy of social behaviour: Changes in cerebral blood flow when people with autistic disorder process facial expressions. *Brain, 123*(Pt 11), 2203–2212.

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., et al. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience, 8*(4), 519–526.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect [slides]*. Palo Alto: Consulting Psychologists Press.

Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., & Evans, A. C. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping, 1*, 214–220.

Frith, C. (2003). What do imaging studies tell us about the neural basis of autism?. *Novartis Found Symposium, 251*, 149–166; discussion 166–176, 281–197.

Frith, C. D., & Frith, U. (1999). Interacting minds—a biological basis. *Science, 286*(5445), 1692–1695.

van der Geest, J. N., Kemner, C., Camfferman, G., Verbaten, M. N., & van Engeland, H. (2002a). Looking at images with human figures: Comparison between autistic and normal children. *Journal of Autism and Developmental Disorders, 32*(2), 69–75.

van der Geest, J. N., Kemner, C., Verbaten, M. N., & van Engeland, H. (2002b). Gaze behavior of children with pervasive developmental disorder toward human faces: A fixation time study. *Journal of Child Psychology and Psychiatry, 43*(5), 669–678.

Gilpin, W. (2002). *Much more laughing and loving with autism*. Arlington: Future Horizons.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41*(17), 2261–2271.

Gosselin, F., & Schyns, P. G. (2004). No troubles with bubbles: A reply to Murray and Gold. *Vision Research*, *44*(5), 471–477; discussion 479–482.

Grandin, T. (1996). *Thinking in pictures: And other reports from my life with autism*. New York: Vintage.

Grelotti, D. J., Gauthier, I., & Schultz, R. T. (2002). Social interest and the development of cortical face specialization: What autism teaches us about face processing. *Developmental Psychobiology, 40*(3), 213–225.

Hadjikhani, N., Joseph, R. M., Snyder, J., Chabris, C. F., Clark, J., Steele, S., et al. (2004). Activation of the fusiform gyrus when individuals with autism spectrum disorder view faces. *Neuroimage, 22*(3), 1141–1150.

Joseph, R. M., & Tanaka, J. (2003). Holistic and part-based face recognition in children with autism. *Journal of Child Psychology and Psychiatry, 44*(4), 529–542.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217–250.

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), 809–816.

Langdell, T. (1978). Recognition of faces: An approach to the study of autism. *Journal of Child Psychology and Psychiatry, 19*(3), 255–268.

LeCouteur, A., Rutter, M., & Lord, C. (1989). Autism diagnostic interview: A standardized investigator-based instrument. *Journal of Autism and Developmental Disorders, 19*, 363–387.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., et al. (1989). Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders, 19*(2), 185–212.

Loveland, K. A., Pearson, D. A., Tunali-Kotoski, B., Ortegon, J., & Gibbs, M. C. (2001). Judgments of social appropriateness by children and adolescents with autism. *Journal of Autism and Developmenal Disorders, 31*(4), 367–376.

Murray, R. F., & Gold, J. M. (2004). Troubles with bubbles. *Vision Research, 44*(5), 461–470.

Ogai, M., Matsumoto, H., Suzuki, K., Ozawa, F., Fukuda, R., Uchiyama, I., et al. (2003). fMRI study of recognition of facial expressions in high-functioning autistic patients. *Neuroreport, 14*(4), 559–563.

Ozonoff, S., & Miller, J. N. (1995). Teaching theory of mind: A new approach to social skills training for individuals with autism. *Journal of Autism and Developmental Disorders, 25*(4), 415–433.

Pedersen, J., Livoir-Petersen, M. F., & Schelde, J. T. (1989). An ethological approach to autism: An analysis of visual behaviour and interpersonal contact in a child versus adult interaction. *Acta Psychiatrica Scandinavica, 80*(4), 346–355.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Pelphrey, K., Adolphs, R., & Morris, J. P. (2004). Neuroanatomical substrates of social cognition dysfunction in autism. *Mental Retardation Developmental Disabilities Research Review, 10*(4), 259–271.

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), 249–261.

Portilla, J., & Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision, 40*, 49–71.

Rogers, S. J. (2000). Interventions that facilitate socialization in children with autism. *Journal of Autism and Developmental Disorders, 30*(5), 399–409.

Rogers, S. J., Hepburn, S. L., Stackhouse, T., & Wehner, E. (2003). Imitation performance in toddlers with autism and those with other developmental disorders. *Journal of Child Psychology and Psychiatry, 44*(5), 763–781.

Schultz, R. T., Grelotti, D. J., Klin, A., Kleinman, J., Van der Gaag, C., Marois, R., et al. (2003). The role of the fusiform face area in social cognition: Implications for the pathobiology of autism. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 358*(1430), 415–427.

Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science, 13*(5), 402–409.

Siegel, B., Vukicevic, J., & Spitzer, R. L. (1990). Using signal detection methodology to revise DSM-III-R: Re-analysis of the DSM-III-R national field trials for autistic disorder. *Journal of Psychiatric Research, 24*(4), 293–311.

Trepagnier, C., Sebrechts, M. M., & Peterson, R. (2002). Atypical face gaze in autism. *Cyberpsychology and Behavior, 5*(3), 213–217.

Triesch, J., Ballard, D. H., Hayhoe, M. M., & Sullivan, B. T. (2003). What you see is what you need. *Journal of Vision, 3*(1), 86–94.

Turatto, M., Angrilli, A., Mazza, V., Umilta, C., & Driver, J. (2002). Looking without seeing the background change: Electrophysiological correlates of change detection versus change blindness. *Cognition, 84*(1), B1–B10.

Volkmar, F. R., Sparrow, S. S., Rende, R. D., & Cohen, D. J. (1989). Facial perception in autism. *Journal of Child Psychology and Psychiatry, 30*(4), 591–598.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing, 13*(4), 600–612.

Weeks, S. J., & Hobson, R. P. (1987). The salience of facial expression for autistic children. *Journal of Child Psychology and Psyhiatry, 28*(1), 137–151.