

Visual Fixation Patterns During Viewing of Naturalistic Social Situations as Predictors of Social Competence in Individuals With Autism

Ami Klin, PhD; Warren Jones, BA; Robert Schultz, PhD; Fred Volkmar, MD; Donald Cohen, MD

Background: Manifestations of core social deficits in autism are more pronounced in everyday settings than in explicit experimental tasks. To bring experimental measures in line with clinical observation, we report a novel method of quantifying atypical strategies of social monitoring in a setting that simulates the demands of daily experience. Enhanced ecological validity was intended to maximize between-group effect sizes and assess the predictive utility of experimental variables relative to outcome measures of social competence.

Methods: While viewing social scenes, eye-tracking technology measured visual fixations in 15 cognitively able males with autism and 15 age-, sex-, and verbal IQ-matched control subjects. We reliably coded fixations on 4 regions: mouth, eyes, body, and objects. Statistical analyses compared fixation time on regions of interest between groups and correlation of fixation time with outcome measures of social competence (ie, standardized

measures of daily social adjustment and degree of autistic social symptoms).

Results: Significant between-group differences were obtained for all 4 regions. The best predictor of autism was reduced eye region fixation time. Fixation on mouths and objects was significantly correlated with social functioning: increased focus on mouths predicted improved social adjustment and less autistic social impairment, whereas more time on objects predicted the opposite relationship.

Conclusions: When viewing naturalistic social situations, individuals with autism demonstrate abnormal patterns of social visual pursuit consistent with reduced salience of eyes and increased salience of mouths, bodies, and objects. Fixation times on mouths and objects but not on eyes are strong predictors of degree of social competence.

Arch Gen Psychiatry. 2002;59:809-816

ADVANCES IN psychological research of the core social deficits in autism have increasingly focused on disruptions in early-emerging skills that seem to derail the processes of socialization.¹ Typically developing infants show preferential attention to social rather than inanimate stimuli,² and they also prefer to focus on the more socially revealing features of the face, such as the eyes rather than the mouth³; in contrast, individuals with autism seem to lack these early social predispositions.⁴⁻⁸ This attenuation to the social world is accompanied by documented abnormalities in face perception, both in face recognition and in identification of facial expressions.⁹⁻¹² In addition, individuals with autism use atypical strategies when performing such tasks, relying on individual pieces of the face rather than on the overall configuration.^{5,8} Alongside these perceptual anomalies, individuals with autism have deficits in conceiving other people's mental states (in having a "theory

that other people have minds and then using this knowledge to predict their social behavior).¹³

Although these findings offer specific hypotheses for the profound impairment in social adjustment exhibited by individuals with autism, there is no performance-based method for directly quantifying the impact of these underlying processes on social functioning in everyday situations. For an individual with autism, the real world is fraught with social challenges and ambiguities that are largely minimized within the narrow parameters of experimental situations. Individuals with autism encounter greater difficulty when trying to spontaneously impose social meaning onto what they see than they do when solving explicit tasks.¹⁴ This is particularly evident in cognitively able individuals, who may use their visual-perceptual or language abilities to scaffold their performance, thus achieving relatively high scores on a given task but doing so in ways that differ from normative strategies.¹⁵ This hypothesis was recently

From the Yale Child Study Center, Yale University, New Haven, Conn.

Table 1. Participant Characterization Data*

| | Autism Group (n = 15) | Control Group (n = 15) | t Value | P Value |
|---|-----------------------------|------------------------------|------------|------------|
| Age, y | 15.4 (7.2) | 17.9 (5.6) | -1.070 | .29 |
| VIQ† | 101.3 (24.9) | 102.5 (20.4) | -0.136 | .89 |
| Vineland socialization standard score‡ | 46.7 (12.7) | ... | ... | ... |
| Vineland socialization age- equivalent score | 4.8 (2.6) | ... | ... | ... |
| ADOS socialization algorithm totals‡ | 9.7 (2.9) | ... | ... | ... |

*Data are given as mean (SD).

†VIQ indicates verbal IQ and is derived from the Wechsler Intelligence Scale for Children, 3rd edition (WISC-III), and the Wechsler Adult Intelligence Scale, 3rd edition (WAIS-III). The WISC-III, WAIS-III, and Vineland have mean = 100 and SD = 15.

‡On Autism Diagnostic Observation Schedule (ADOS)–3 and ADOS-4, the cut-off point for autism is 6 on the socialization algorithm items (maximum score is 14).

substantiated in a neurofunctional study¹⁶ of face perception in autism in which adequate task performance was accompanied by abnormal ventral temporal cortical activity, which in turn suggested that participants had “treated” faces as objects.

The development of more naturalistic paradigms that capitalize on the challenging nature of open-ended social scenarios for individuals with autism could result in several advantages. First, the effect size of between-group findings may increase, bringing experimentally measured abnormalities in autism to a level more commensurate with clinical observations of social impairment. Second, by placing a performance-based method into the context of naturalistic social functioning, there is increased likelihood that the experimental method will better predict level of social competence. Abnormal performance on social tasks in autism is typically reported as a group finding relative to a control sample, with little attempt to use the given measure as a predictor of level of social adjustment or of level of social impairment. The lack of quantifiable indices of social competence that could define a spectrum of social outcomes, from normality to varying manifestations of autism, hinders genetic and neurofunctional research, which partially rely on such indices for interpretation of heritability and neuroimaging data.^{17,18}

To create an experimental paradigm to measure social functioning in a context that more closely resembles social demands in naturalistic situations, we used eye-tracking technology to study spontaneous viewing patterns of cognitively able individuals with autism and age- and verbal IQ–matched control subjects. Previous work^{19,20} has used static images to assess aspects of face scanning; to our knowledge, this is the first study of eye tracking as a dynamic phenomenon. The paradigm involves viewing digitized videotape clips of complex social situations while wearing a noninvasive eye-tracking device that superimposes the viewer's point of regard onto the viewed scenes. The resultant videotape can then be coded for patterns of visual fixation that can be measured in terms of percentage of viewing time spent on different aspects of the social scenes. For this exploratory study, several lines of research guided

our decision as to the scheme to use for coding fixations. First, we were interested in the relative salience of major components of the viewed scenes.^{4,8,14,15} Thus, we divided the total on-screen area into face, body, and object regions. Second, given the research results suggesting that individuals with autism have difficulty interpreting social information conveyed through the eyes,^{6,7} and that they focus preferentially on mouths when performing face perception tasks,⁵ we subdivided the face into an eye region and a mouth region. The resultant coding scheme, therefore, consisted of 4 clearly delineated viewing areas of interest: mouth, eye, body, and objects (see the “Coding Procedures” subsection for further details). We were also interested in exploring the relationship between patterns of visual fixation and outcome measures of social competence. We operationalized social competence in terms of level of real-life social adjustment (as defined by scores on the socialization domain of the Vineland Adaptive Behavior Scales, Expanded Edition [VABS-E]²¹) and in terms of degree of autistic social symptoms (as defined by the social domain of the Autism Diagnostic Observation Schedule [ADOS]²²). Whereas the VABS-E provides a measure of social ability (higher scores mean higher levels of social adjustment), the ADOS provides a measure of social impairment (higher scores mean higher levels of autistic social behaviors).

On the basis of the extant literature, we expected that individuals with autism would preferentially focus on the mouth region rather than on the eye region and on objects relative to controls. In addition, given the normative pattern of preferentially using information gathered from the eyes to understand others, we also predicted that within the autistic group, higher percentage of viewing time on the eye region would be positively correlated with level of social adjustment and negatively correlated with level of social impairment.

PARTICIPANTS AND METHODS

PARTICIPANTS

Fifteen male adolescents and young adults with autism were recruited through a large, federally funded research project on the neurobiology of autism carried out in the developmental disabilities section of the Yale Child Study Center. This project includes a 3-day protocol consisting of extensive diagnostic, neuropsychological, neuroimaging, and genetic studies. Before participation, all individuals or their legal guardians supplied written informed consent. The protocol was approved by the Human Investigations Committee of the Yale University School of Medicine. Diagnoses were assigned on the basis of parental interview (Autism Diagnostic Interview–Revised²³) and direct observations of participants' social and communicative behaviors (ADOS²²). All participants with autism met DSM-IV²⁴ diagnostic criteria for autistic disorder as operationalized through standardized algorithms derived from the Autism Diagnostic Interview–Revised and the ADOS. Intellectual level was measured using the Wechsler Intelligence Scale for Children²⁵ or the Wechsler Adult Intelligence Scale.²⁶ All 15 participants with autism were cognitively able and yet severely impaired in their social functioning, as measured using the VABS-E¹⁹: a discrepancy of more than 3.5 SD was obtained between their verbal IQ and standard scores on the socialization domain of the VABS-E (see **Table 1** for a summary of participant characterization data). Participants with autism were individually matched

for chronological age and verbal IQ with a comparison group of 15 adolescents and young adults recruited from the community and screened for history of major neurological or psychiatric illness. Given the typical variability in IQ profile in individuals with autism and the resultant complications when making decisions on matching procedures in studies of social functioning,^{27,28} we adopted what is thought to be the most stringent approach to matching on the basis of verbal IQ rather than full-scale IQ. None of the participants had visual acuity deficits uncorrectable with eyeglasses.

DIGITIZED VIDEOTAPE CLIPS OF COMPLEX SOCIAL SITUATIONS

All participants watched 5 digitized clips from the 1967 film version of Edward Albee's "Who's Afraid of Virginia Woolf?" This movie was chosen because it displays the intense interaction of 4 protagonists involved in a content-rich social situation likely to maximize viewers' monitoring of each person's socially expressive actions as well as those characters' reactions to the actions of others. The demanding social complexity in the movie mirrors complicated social situations that individuals with autism may encounter in everyday settings, such as a high-school cafeteria. This movie is also depleted of nonessential objects and events that might distract a viewer's attention from the social action. The clips ranged in length from 30 to 60 seconds depending on the content and duration of each chosen scene. The clips were separated from one another by 5 seconds of blank screen.

APPARATUS

Eye tracking was accomplished using a dark pupil-corneal reflection video-oculography technique and hardware and software created by ISCAN Inc (Burlington, Mass). The system was head mounted and used a novel target-tracking method that enabled highly accurate eye tracking without having to restrain the participant's head (accuracy within $\pm 0.3^\circ$ over a $\pm 20^\circ$ horizontal and vertical range). The eye-tracking video equipment was mounted unobtrusively on the bill of a baseball cap. The participant's left eye was illuminated by a small, nonharmful infrared light-emitting diode. The movements of this eye were filmed using miniature imaging optics and a dichroic mirror. To obtain a frontal image of the eye without interfering with the participant's view, the eye-imaging camera was mounted in the bill above, facing downward and aimed at the angled dichroic mirror. The mirror, positioned in front of and below the eye, acted as a bandpass filter, reflecting infrared wavelengths (and thereby reflecting a frontal image of the participant's left eye) while transmitting wavelengths of the visible spectrum (allowing the participant a clear field of view). Directly underneath this mirror and in line with the eye-imaging optics was a miniature camera capped by a red additive filter. This camera filmed the scene in front of each participant. As the eye camera tracked movements of the pupil and corneal reflection, the scene camera recorded images of the participant's field of view. In a rectangular configuration around the computer screen were 5 red light-emitting diodes (1 at each of the screen's corners and 1 at the midpoint between the 2 upper corners) that provided reference points by which image-processing hardware (using data from the scene camera, aided by the red color filter mentioned previously) could identify the position and orientation of the "scene plane" even within a shifting field of view. In this way, the coordinates of the plane of the computer screen could be continuously tracked (and then calibrated with respect to the coordinates of the point of regard) despite any head or body movements. Eye-tracking data were fed from the cameras through ISCAN hardware at a rate of 60 samples per sec-



Figure 1. Representative still frame used for coding of visual fixation patterns. The data of 2 participants are superimposed on a single frame for the purpose of illustration; the focus of the view with autism is marked in green and that of the typically developing viewer (control) is in yellow. Coordinate data shown in the inset correspond to the typically developing viewer's point of regard.

ond and were recorded to videotape at the standard rate of 30 frames per second.

EXPERIMENTAL SETTING AND PROCEDURES

Participants sat in a comfortable armchair, 63.5 cm from a 48.3-cm computer screen mounted flush within a black wooden panel. The eye-tracking baseball cap was adjusted for participant comfort and clarity of view, and a brief calibration routine followed (consisting of having each participant look at a series of 5 points). Lights in the room were dimmed so that only images displayed on the screen could be easily seen, and the audio component was played through a set of concealed speakers. Data recording began after each participant reported an adequate level of comfort, an unobstructed view of the screen, and a clearly audible soundtrack.

CODING PROCEDURES

Following each recording session, the videotape data were digitized and archived on computer. Five videotape clips lasting a total of 2 minutes 42 seconds were coded for each participant. Each frame of video was coded (30 frames per second, 4860 frames per participant). **Figure 1** exemplifies a still image used for coding. For this illustration, the points of regard of 2 participants (1 with autism and 1 control) are superimposed onto a single frame. In practice, only one participant's data were coded at a time by 2 raters who were blind to the diagnosis and identity of each participant. Kappa coefficients²⁹ and percentages of agreement were calculated to assess interrater reliability, with κ of 0.82 and agreement of 87.2%, indicating "excellent agreement."³⁰ Each rater also recoded approximately 20% of the scenes chosen at random after at least 2 weeks. Test-retest reliability results were $\kappa=0.91$ (agreement, 95.2%) and $\kappa=0.86$ (agreement, 91.3%) for the 2 coders, indicating excellent levels of coding stability.

For each frame, the location of a participant's point of regard was coded as follows: a designation of "0" was assigned when the participant focused on the mouth region of the face, "1" was recorded when he focused on the eye region of the face, "2" was given when he focused on the body, and "3" was recorded when he focused on an object. In the event that a participant's point of regard was directly on the border between any 2 regions of interest, the 3 frames before and the 3 frames immediately after the frame in question were evaluated. The frame was then coded in accordance with the region into which the point of regard moved during the 7-frame sequence. To minimize such circum-

Table 2. Percentage of Viewing Time Spent Focused on Mouth, Eyes, Body, and Object Regions*

| Region | Autism Group (n = 15) | Control Group (n = 15) | t Value | P Value |
|--------|-----------------------|------------------------|---------|---------|
| Mouth | 41.2 (15.0) | 21.2 (12.1) | 4.026 | <.001 |
| Eyes | 24.6 (8.1) | 65.4 (12.8) | -10.455 | <.001 |
| Body | 24.6 (12.4) | 9.7 (5.7) | 4.226 | <.001 |
| Object | 9.6 (6.5) | 3.7 (2.4) | 3.286 | <.003 |

*Data are given as mean (SD).

stances, region-of-interest resolution was maximized: only those shots in which the on-screen dimensions of the character's head were greater than or equal to 5° of the participant's field of view could be coded (this was the basis for selecting the 2 minutes 42 seconds of codable videotape data from the 5 clips shown to each participant; long-range shots were excluded). This criterion was used to ensure validity of the eye-tracking data coding, particularly in relation to coding of the eye vs the mouth region (if an on-screen head is smaller than 5° of the visual field, the system's margin of error is large enough to potentially cause notable discrepancies in the coding of facial feature preference). Finally, several scenes were discarded because there were no actors facing the camera or because the actors, the camera action, or both moved too rapidly for meaningful analysis of preferential viewing patterns.

Some frames that were coded could not be coded as fixation data. These included frames in which a participant blinked or made a saccadic shift. Blinks were defined by a loss of point-of-regard coordinate data (which was superimposed along with time code at the bottom of each picture) and then further verified by an inset videotaped image of the participant's eye (in such a case revealing closed lids). Initiation and completion of saccadic movements were defined by a rotational velocity threshold of 30° per second,³¹ which translated in our experimental setting into a linear velocity of 15 or more pixels per frame (ie, $d \geq 15$ pixels, where $d = \sqrt{(x^2) + (y^2)}$, over the course of 1 frame [33 milliseconds]). Saccadic movements were also confirmed by the inset eye image. In addition, some frames could not be coded for the following reasons: off-screen fixations (watches, hands, etc), rubbing of eyes, and obstruction of the scene view camera (hand in front of face, nose rubbing, etc). These frames, recorded as "no data," amounted to a mean (SD) 5.2% (6.6%) of total time for controls and 15.6% (15.9%) of time for viewers with autism ($P = .03$). Although this significant difference could be related to decreased concentration and increased distractibility in participants with autism, we cannot infer any specific conclusions from this statistical difference because other, non-attention-related factors could also be at play, including, for example, rubbing the eyes or sneezing. Future studies will need to pay more attention to coding schema for characterization of lost data.

STATISTICAL ANALYSES

T tests with Bonferroni corrections for number of between-group comparisons were used to test group differences in the percentage of total viewing time spent on fixations on mouth, eye, body, and object regions of the scene images. Pearson correlations were used to explore the degree of the relationship between fixation patterns and the 2 measures of social competence—level of social adjustment (operationalized as the age-equivalent score on the socialization domain of the VABS-E) and degree of autistic social impairment (operationalized as the total algorithm score on the social section of the ADOS).

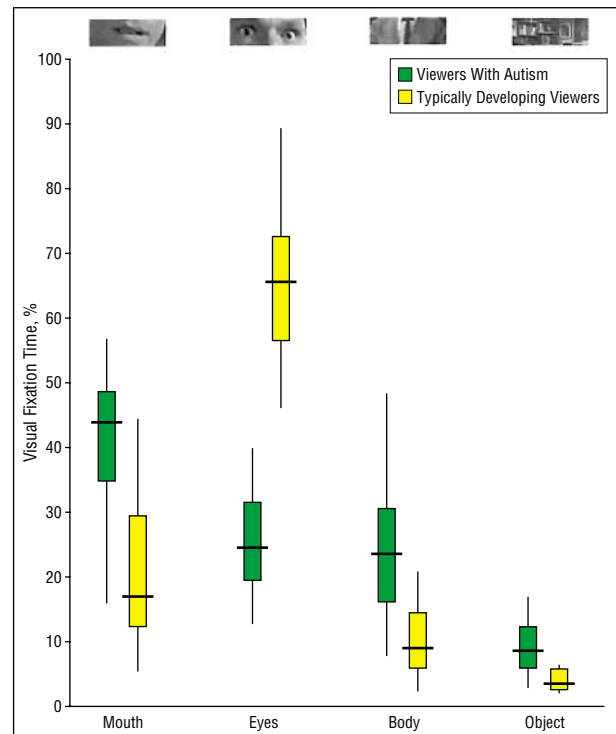


Figure 2. Box plot comparison of visual fixation time on mouth, eyes, body, and object regions for 15 viewers with autism and 15 typically developing viewers (controls). The upper and lower boundaries of the standard boxplots are at the 25th and 75th percentiles. The horizontal line across the box marks the median of the distribution, and the vertical lines below and above the box extend to the minimum and maximum, respectively.

RESULTS

Consistent with our predictions, individuals with autism focused 2 times more on the mouth region, 2 times less on the eye region, 2 times more on the body region, and 2 times more on the object region relative to age- and verbal IQ-matched controls (**Table 2**). Effect size was greatest for fixation on the eye region, making it the best predictor of group membership ($d = 3.19$). There were no significant correlations between any of the measures of fixation time and chronological age or verbal IQ in the 2 groups except for a positive significant correlation between percentage of fixation time on the mouth region and age in the group with autism ($r = 0.62$; $P = .01$). When one outlier (a participant with autism aged 38 years) was removed from the sample, this correlation was no longer significant. None of the other correlations were altered by excluding this person. As shown in **Figure 2**, there was little overlap in the measures of fixation time across the 2 groups, although there was some variability within the groups, except for fixation on objects in the control group.

We next explored the association between fixation time measures and measures of social competence. Contrary to our expectation, fixation time on the eye region was not associated with either social adaptation (VABS-E socialization scores) or social disability (ADOS social scores) ($r = -0.20$ and $r = 0.14$, respectively). In contrast, fixation times on the mouth region (**Figure 3**) and on the object region (**Figure 4**) were strong predictors of

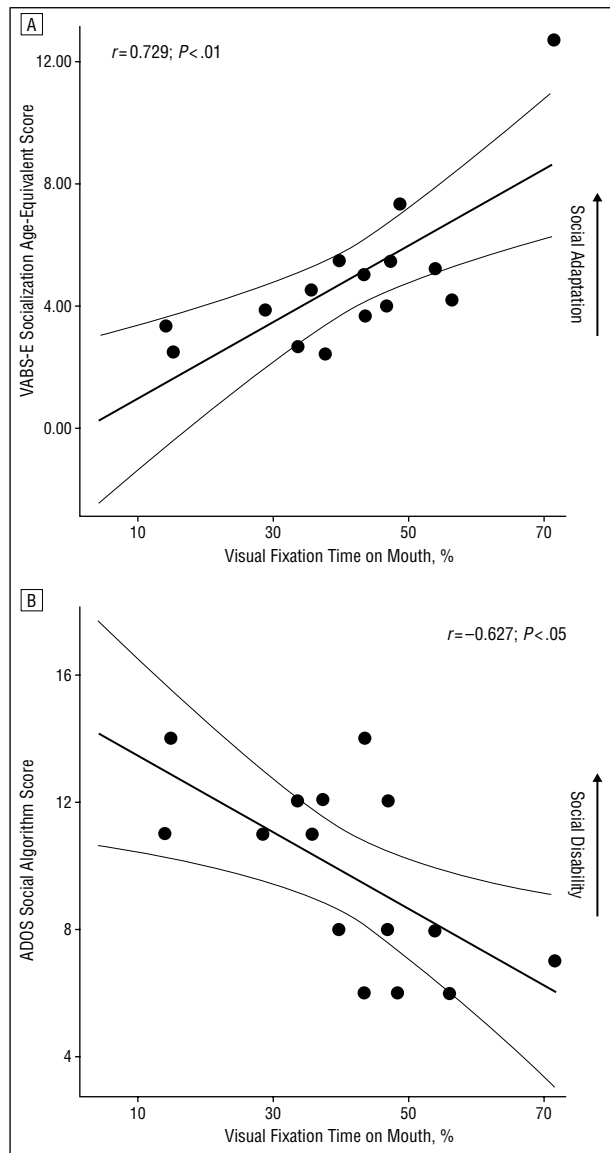


Figure 3. Correlation of mouth fixation time and outcome measures of social adaptation (A) and social disability (B) in 15 viewers with autism. VABS-E indicates Vineland Adaptive Behavior Scales, Expanded Edition; ADOS, Autism Diagnostic Observation Schedule.

social competence, albeit in different directions. Fixation time on the mouth region was associated with greater social adaptation (ie, more socially able) and lower autistic social impairment (ie, less socially disabled). Going in the opposite direction, fixation time on the object region was associated with lower social adaptation and greater autistic social impairment. Fixation time on the body region followed the same trend as for the object region, but the correlations were not significant ($r=-0.49$ and $r=0.34$ for social adaptation and social disability, respectively). When the outlier (a participant with autism aged 38 years) was excluded from correlational analyses, none of the measures of association were altered except the positive correlation between fixation time on objects and the measure of social adaptation, which was previously significant at $P<.10$ and was now significant at $P<.05$.

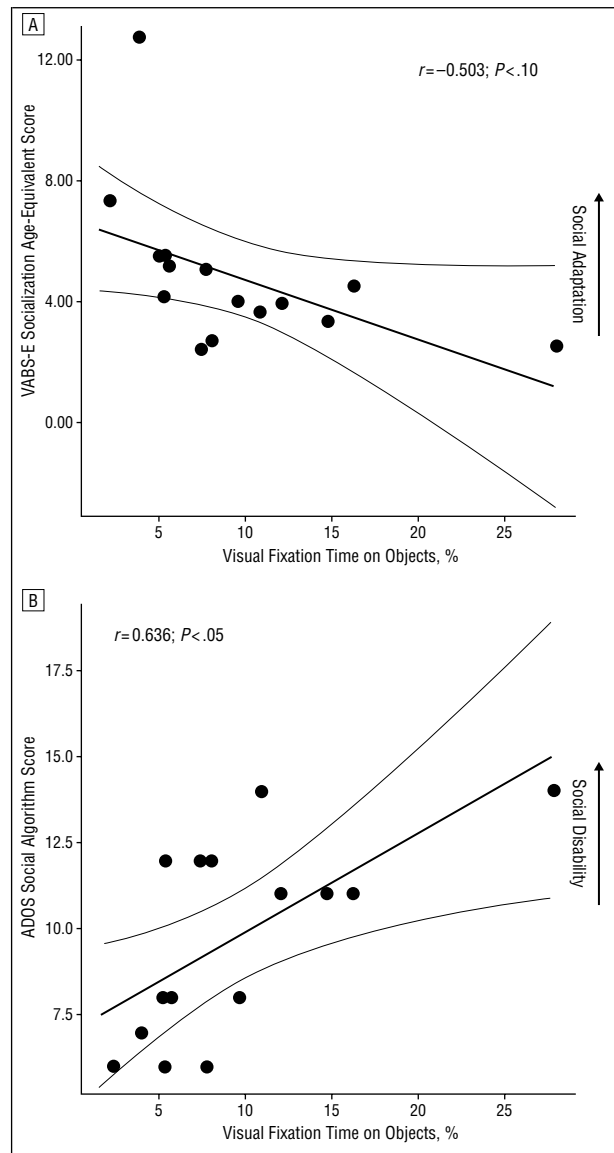


Figure 4. Correlation of object fixation time and outcome measures of social adaptation (A) and social disability (B) in 15 viewers with autism. VABS-E indicates Vineland Adaptive Behavior Scales, Expanded Edition; ADOS, Autism Diagnostic Observation Schedule.

COMMENT

We found significant differences in percentages of visual fixation time on mouth, eye, body, and object regions when viewing naturalistic social situations among cognitively able adolescents and young adults with autism relative to age- and verbal IQ-matched controls. The best predictor of group membership was percentage of fixation time on the eye region: the control group visually fixated on the eye region 2 times more than did the group with autism, and there was no overlap in the distribution of results. However, within the group with autism, this variable was unrelated to outcome measures of social competence. In contrast, percentages of fixation time on the mouth and object regions were strong predictors of social competence measures, with higher mouth fixation time associated with higher levels of social adaptation and lower levels of autistic social impair-

ment, and object fixation time associations going in the opposite direction.

This study shows the utility of developing performance-based paradigms capable of quantifying social functioning under more naturalistic conditions. The effect sizes obtained for between-group comparisons were markedly larger relative to experimental studies³² of social functioning in autism and more comparable to the few studies¹⁴ available that focused on spontaneous responses rather than explicit problem-solving tasks. Similarly, the enhanced ecological validity aimed for in the present study also led to stronger associations between performance-based experimental measures and outcome measures of social competence, an important goal to be achieved in the search for dimensional endophenotypes for genetic research.¹⁷ More generally, the paradigm presented herein provides a valuable window into the ways individuals with autism search for meaning when confronted with complex social situations.

Given that this is the first study of its kind, comparisons with other studies can only be indirect. Nevertheless, the small number of available studies is generally supportive of our findings. For example, some studies have suggested increased reliance on mouths rather than eyes when individuals with autism are required to perform face perception tasks,^{5,6} great difficulty in “reading” the meaning of eye expressions,⁷ and increased orientation to objects relative to people.^{4,8} A recent study by Joseph³³ addressed the differential reliance on eyes and mouths in autism more directly than previous studies. When 10- to 12-year-old, cognitively able individuals with autism were asked to perform a match-to-sample face recognition task in which faces varied as a function of eye or mouth changes, they exhibited a selective advantage for mouths.

Although the present study was not designed to explore the factors underlying visual fixation patterns in autism, the strong predictive association between mouth and object, but not eye fixations and outcome measures of social competence, suggests that specific hypotheses need to be pursued in future, more-refined studies. The finding that a higher percentage of mouth fixation time predicted a higher level of social competence is as intriguing as it is counterintuitive, and, therefore, it deserves special attention. Given the well-known association between level of verbal skills and better outcome in autism,³⁴ it is possible that the participants in our study were focusing on mouths because that is where speech comes from. By concentrating their efforts on something that they can understand, they might attain better understanding of social situations. Nevertheless, such a compensatory strategy is not without its limitations given that the meaning of language is often modified by non-verbal social cues such as eye expressions. The fact that percentage of time focused on eyes was unrelated to measures of social competence suggests the possibility that for individuals with autism, looking at eyes does not accrue considerable advantages in their efforts to understand social situations, or, in other words, that the eyes are not meaningful to them.^{7,34}

The hypothesis that increased fixation on mouths results from concentration on speech needs to be fur-

ther refined on the basis of what we know about visual attention to facial regions in speech perception tasks.³⁵ Whereas audiovisual speech perception (like the stimuli in the present study) draws gaze toward the talker's eyes,³⁶ speech reading (ie, seeing someone speak with no accompanying audio) is more likely to draw gaze toward the talker's mouth.³⁷ However, visual fixations in speech reading depend on the nature of the task. If participants are asked to perform a segmental speech perception task (ie, to guess what the talker is saying), the viewer is more likely to focus on the mouth region.³⁶⁻³⁸ However, if participants are asked to perform a prosodic speech perception task (ie, to guess the inflection of the talker's voice, eg, sad, happy, or mad), then the viewer is more likely to focus on the upper region of the face (eyes and eyebrows).³⁷ On the basis of this literature, our visual fixation results could suggest that our participants were not only focusing on the verbal content of speech but were also ignoring paralinguistic cues such as prosody,³⁹ which are usually essential to understanding nonliteral aspects of social situations such as intentions and attitudes.⁴⁰ In other words, our findings could be related to the over-reliance of individuals with autism on the literal aspects of speech at the expense of intonational cues associated with social meaning.⁴¹ By pairing speech perception tasks with eye-tracking measures,³⁷ these relationships are likely to be clarified.

Another line of inquiry for future studies relates to the possibility that increased fixation on mouths and decreased fixation on eyes indicate that individuals with autism may acquire a degree of perceptual expertise on mouths but not on eyes. Findings of the previously described study by Joseph³³ seem to point in this direction. When the face recognition task was presented in the form of inverted faces, children with autism did not show the typical decrement in performance⁴² (the so-called inversion effect⁴³) in the eye condition, but did so in the mouth condition. Given that an increase in the magnitude of the inversion effect and associated transition from feature-based to configural processing marks the development of perceptual expertise relative to a class of objects,⁴⁴⁻⁴⁷ the findings of Joseph raise the possibility that children with autism have expertise when recognizing faces that vary in mouth features but not when they vary in eye features. Future studies that could manipulate these variables more explicitly in still (ie, photographs) and dynamic (ie, videotape) stimuli might clarify these issues. Also, given the recent surge of neuroimaging research on the brain circuitry involved in the acquisition of perceptual expertise,⁴⁷ the coregistration of neurofunctional and eye-tracking data might prove to be particularly useful in elucidating the unusual patterns of behavioral and brain functioning in regard to face perception in autism.^{16,18,48}

A final line of inquiry suggested by our results concerns the need to explore the association between more fixation time on objects and decreased social competence in participants with autism. This finding is consistent with the notion that focus on objects means reduced salience of social stimuli^{4,15} and, therefore, decreased likelihood of understanding the social situation. The percentage of fixation time on objects was small relative to

fixation time on the face (mouth and eyes). This is not surprising given that our videotape clips were deliberately chosen to minimize inanimate distractions. Future studies may maximize the predictive utility of the association between fixation on objects and outcome measures of social competence by manipulating the prominence of inanimate stimuli (eg, by making objects move or by using objects known to attract the attention of individuals with autism). Clinical observations have shown the devastating impact of object-rich environments on the ability of children with autism to focus on socially meaningful aspects of educational environments.⁴⁹

The present study has several limitations. First, although close to three quarters of all individuals with autism have a degree of mental retardation,⁵⁰ we studied more cognitively able individuals with autism. We do not know if these results will extend to participants with lower IQs. Neither do we know what trends to expect in younger children with autism or in individuals with milder manifestations of the condition (eg, Asperger syndrome and Pervasive Developmental Disorder–Not Otherwise Specified [also called “atypical autism”]). Second, our design does not rule out possible contributions from abnormal functioning in areas other than social visual pursuit (eg, underlying attentional or perceptual abnormalities).^{51,52} The use of more traditional eye-tracking protocols focused on the integrity of brain mechanisms associated with eye movement⁵³ alongside our novel application of this technology will be necessary to rule out explanations other than the ones explored in this study. Third, although the measures reported in this study were informative, they are unlikely to be the most sensitive indicators of abnormalities in social visual pursuit in autism. In a case study reported elsewhere,⁵⁴ we illustrate a series of social visual tracking phenomena that seem to capture the abnormalities in autism in a much more dynamic and stark fashion than summaries of visual fixation times. The moment-by-moment visual traces left behind by the saccadic movements and fixations of individuals with autism illustrate more vividly their atypical attempts to create social meaning out of what they see. Perhaps paradigms capable of quantifying the subtleties of such data will soon emerge from these as-yet only heuristic approaches.

Submitted for publication June 14, 2001; final revision received October 10, 2001; accepted October 24, 2001.

This work was partly supported by grants P01 HD 03008 and P01 HD/DC35482 from the National Institute of Child Health and Human Development, Bethesda, Md, and by a grant from the National Alliance for Autism Research, Princeton, NJ.

This study was presented in part at the annual meeting of the National Institute of Child Health and Human Development, Collaborative Projects of Excellence in Autism, New Haven, Conn, May 2, 2001.

We thank the participants and their families; Elizabeth Levitan, BA, Tammy Babitz, MA, Sanno Zack, BA, Catalina Hooper, BA, Amy Augustyn, BA, and Christopher Abildgaard, BA, for their assistance with this study; and Katarzyna Chawarska, PhD, for her contribution to interpretation of findings. We also thank Warner

Bros, Elizabeth Taylor, George Segal, and the estates of Sandy Dennis and Richard Burton for permission to use their images.

This article is dedicated to the memory of our mentor, collaborator, and colleague, Donald Cohen, MD, whose legacy embodies the best in clinical services, public advocacy, clinical science, bioethics, and mentoring.

Corresponding author and reprints: Ami Klin, PhD, Yale Child Study Center, 230 S Frontage Rd, New Haven, CT 06520 (e-mail: ami.klin@yale.edu).

REFERENCES

1. Waterhouse L, Fein D, Modahl C. Neurofunctional mechanisms in autism. *Psychol Rev*. 1996;103:457-489.
2. Spelke ES, Phillips A, Woodward AL. Infants' knowledge of object motion and human action. In: Sperber D, Premack D, Premack AJ, eds. *Causal Cognition: A Multidisciplinary Debate*. Oxford, England: Oxford University Press; 1995:44-78.
3. Haith MM, Bergman T, Moore MJ. Eye contact and face scanning in early infancy. *Science*. 1979;198:853-855.
4. Dawson G, Meltzoff AN, Osterling J, Rinaldi J, Brown E. Children with autism fail to orient to naturally occurring social stimuli. *J Autism Dev Disord*. 1998;28:479-485.
5. Langdell T. Recognition of faces: an approach to the study of autism. *J Child Psychol Psychiatry*. 1978;19:255-268.
6. Baron-Cohen S, Wheelwright S, Jolliffe T. Is there a “language of the eyes”? evidence from normal adults, and adults with autism or Asperger syndrome. *Vis Cogn*. 1997;4:311-331.
7. Baron-Cohen S, Jolliffe T, Mortimore C, Robertson M. Another advanced test of theory of mind: evidence from very high functioning adults with autism or Asperger syndrome. *J Child Psychol Psychiatry*. 1997;38:813-822.
8. Hobson RP, Ouston J, Lee A. What's in a face? the case of autism. *Br J Psychol*. 1988;79:441-453.
9. Klin A, Sparrow SS, de Bildt A, Cicchetti DV, Cohen DJ, Volkmar FR. A normed study of face recognition in autism and related disorders. *J Autism Dev Disord*. 1999;29:497-507.
10. Tantam D, Monaghan L, Nicholson H, Stirling J. Autistic children's ability to interpret faces: a research note. *J Child Psychol Psychiatry*. 1989;30:623-630.
11. Boucher J, Lewis V. Unfamiliar face recognition in relatively able autistic children. *J Child Psychol Psychiatry*. 1992;33:843-859.
12. Weeks S, Hobson R. The salience of facial expression for autistic children. *J Child Psychol Psychiatry*. 1987;28:137-151.
13. Baron-Cohen S, Tager-Flusberg H, Cohen D, eds. *Understanding Other Minds: Perspectives From Developmental Neuroscience*. 2nd ed. Oxford, England: Oxford University Press; 2000.
14. Klin A. Attributing social meaning to ambiguous visual stimuli in higher functioning autism and Asperger syndrome: the social attribution task. *J Child Psychol Psychiatry*. 2000;41:831-846.
15. Klin A, Schultz R, Cohen D. Theory of mind in action: developmental perspectives on social neuroscience. In: Baron-Cohen S, Tager-Flusberg H, Cohen D, eds. *Understanding Other Minds: Perspectives From Developmental Neuroscience*. 2nd ed. Oxford, England: Oxford University Press; 2000:357-388.
16. Schultz RT, Gauthier I, Klin A, Fulbright R, Anderson A, Volkmar FR, Skudlarski P, Lacadie C, Cohen DJ, Gore JC. Abnormal ventral temporal cortical activity among individuals with autism and Asperger syndrome during face discrimination. *Arch Gen Psychiatry*. 2000;57:331-340.
17. Rutter M, Silberg J, O'Connor T, Simonoff E. Genetics and child psychiatry, I: advances in quantitative and molecular genetics. *J Child Psychol Psychiatry*. 1999;40:3-18.
18. Schultz RT, Romanski LM, Tsatsanis KD. Neurofunctional models of autistic disorder and Asperger syndrome: clues from neuroimaging. In: Klin A, Volkmar F, Sparrow S, eds. *Asperger Syndrome*. New York, NY: Guilford Publications; 2000:172-209.
19. Yarbus AL. *Eye Movements and Vision*. New York, NY: Plenum Press; 1967.
20. Walker-Smith GJ, Gale AG, Findlay JM. Eye movement strategies in face perception. *Perception*. 1977;6:313-326.
21. Sparrow SS, Balla D, Cicchetti D. *Vineland Adaptive Behavior Scales, Expanded Edition*. Circle Pines, Minn: American Guidance Service; 1984.
22. Lord C, Rutter M, DiLavore P, Risi S. *Autism Diagnostic Observation Schedule (ADOS)*. Los Angeles, Calif: Western Psychological Services; 1999.
23. Lord C, Rutter M, LeCouteur A. Autism Diagnostic Interview–Revised: a revised

- version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord*. 1994;24:659-685.
24. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. Washington, DC: American Psychiatric Association; 1994.
 25. Wechsler D. *Wechsler Intelligence Scale for Children*. 3rd ed. San Antonio, Tex: Psychological Corp; 1991.
 26. Wechsler D. *Wechsler Adult Intelligence Scale*. 3rd ed. San Antonio, Tex: Psychological Corp; 1997.
 27. Klin A, Carter A, Volkmar FR, Cohen DJ, Marans WD, Sparrow SS. Assessment issues in children with autism. In: Cohen DJ, Volkmar FR, eds. *Handbook of Autism and Pervasive Developmental Disorders*. 2nd ed. New York, NY: John Wiley & Sons Inc; 1997:411-418.
 28. Hobson RP. Methodological issues for experiments on autistic individuals' perception and understanding of emotion. *J Child Psychol Psychiatry*. 1991;32:1135-1158.
 29. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Measure*. 1960;20:37-46.
 30. Cicchetti DV, Sparrow SS. Developing criteria for establishing inter-rater reliability of specific items in a given inventory. *Am J Ment Defic*. 1981;86:127-137.
 31. Leigh RJ, Zee DS. *The Neurology of Eye Movements*. 3rd ed. Oxford, England: Oxford University Press; 1999:94.
 32. Yirmiya N, Erel O, Shaked M, Solomonica-Levi D. Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. *Psychol Bull*. 1998;124:283-307.
 33. Joseph RM. Face recognition processes in typically developing children and children with autism. Paper presented at: Society for Research in Child Development Biennial Meeting; April 21, 2001; Minneapolis, Minn.
 34. Lord C, Paul R. Language and communication in autism. In: Cohen DJ, Volkmar FR, eds. *Handbook of Autism and Pervasive Developmental Disorders*. New York, NY: John Wiley & Sons Inc; 1997:195-225.
 35. Storke DG, Hennecke ME. Human speechreading: psychology and cognition. In: *Speechreading in Humans and Machines: Models, Systems, and Applications*. Berlin, Germany: Springer-Verlag; 1996:525-535.
 36. Vatikiotis-Bateson E, Eigsti I-M, Yano S, Munhall KG. Eye movement of perceivers during audiovisual speech perception. *Percept Psychophys*. 1998;60:926-940.
 37. Lansing CR, McConkie GW. Attention to facial regions in segmental and prosodic visual speech perception tasks. *J Speech Lang Hear Res*. 1999;42:526-539.
 38. Preminger JE, Lin H-B, Payen M, Levitt H. Selective visual masking in speechreading. *J Speech Lang Hear Res*. 1998;41:564-575.
 39. Scherer KR. Vocal affect expression: a review and a model for future research. *Psychol Bull*. 1986;99:143-165.
 40. Bolinger DL. *Intonation and Its Parts*. Stanford, Calif: Stanford University Press; 1986.
 41. Landa R. Social language use in Asperger syndrome and high-functioning autism. In: Klin A, Volkmar FR, Sparrow SS, eds. *Asperger Syndrome*. New York, NY: Guilford Publications; 2000:125-155.
 42. Tanaka JW, Kay JB, Grinnell E, Stansfield B, Szechter L. Face recognition in young children: when the whole is greater than the sum of its parts. *Vis Cogn*. 1998;5:479-496.
 43. Valentine T. Upside-down faces: a review of the effect of inversion upon face recognition. *Br J Psychol*. 1988;79:471-491.
 44. Gauthier I, Tarr MJ. Becoming a "Greeble" expert: exploring mechanisms for face recognition. *Vision Res*. 1997;37:1673-1681.
 45. Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC. Activation of the middle fusiform "face area" increases with expertise in recognizing novel objects. *Nat Neurosci*. 1999;2:568-573.
 46. Diamond R, Carey S. Why faces are and are not special: an effect of expertise. *J Exp Psychol Gen*. 1986;115:107-117.
 47. Gauthier I, Skudlarski P, Gore JC, Anderson AW. Expertise for cars and birds recruits brain areas involved in face recognition. *Nat Neurosci*. 2000;3:191-197.
 48. Marcus DJ, Nelson CA. Neural bases and development of face recognition in autism. *CNS Spectrums*. 2001;6:36-59.
 49. Olley JG, Reeve C. Issues of curriculum and classroom structure. In: Cohen DJ, Volkmar FR, eds. *Handbook of Autism and Pervasive Developmental Disorders*. New York, NY: John Wiley & Sons Inc; 1997:484-508.
 50. Fombonne E. The epidemiology of autism: a review. *Psychol Med*. 1999;29:769-786.
 51. Harris NS, Courchesne E, Townsend J, Carper RA, Lord C. Neuroanatomic contributions to slowed orienting of attention in children with autism. *Cognit Brain Res*. 1999;8:61-71.
 52. Courchesne E, Townsend J, Akshoomoff NA, Saitoh O, Yeung-Courchesne R, Lincoln AJ, James H, Hass R, Schreibman L, Lau L. Impairments in shifting attention in autistic and cerebellar patients. *Behav Neurosci*. 1994;108:1-18.
 53. Minshew NJ, Luna B, Sweeney JA. Oculomotor evidence for neocortical systems but not cerebellar dysfunction in autism. *Neurology*. 1999;52:917-922.
 54. Klin A, Jones W, Schultz R, Volkmar FR, Cohen DJ. Defining and quantifying the social phenotype in autism. *Am J Psychiatry*. In press.