Development of an Objective Autism Risk Index Using Remote Eye Tracking

Thomas W. Frazier, PhD, Eric W. Klingemier, BA, Mary Beukemann, Leslie Speer, PhD, NCSP, Leslie Markowitz, PsyD, Sumit Parikh, MD, Steven Wexberg, MD, Kimberly Giuliano, MD, Elaine Schulte, MD, MPH, Carol Delahunty, MD, Veena Ahuja, MD, Charis Eng, MD, PhD, Michael J. Manos, PhD, Antonio Y. Hardan, MD, Eric A. Youngstrom, PhD, Mark S. Strauss, PhD

Objective: Abnormal eye gaze is a hallmark characteristic of autism spectrum disorder (ASD), and numerous studies have identified abnormal attention patterns in ASD. The primary aim of the present study was to create an objective, eye tracking-based autism risk index.

Method: In initial and replication studies, children were recruited after referral for comprehensive multidisciplinary evaluation of ASD and subsequently grouped by clinical consensus diagnosis (ASD n = 25/15, non-ASD n = 20/19 for initial/replication samples). Remote eye tracking was blinded to diagnosis and included multiple stimuli. Dwell times were recorded to each a priori–defined region of interest (ROI) and averaged across ROIs to create an autism risk index. Receiver operating characteristic curve analyses examined classification accuracy. Correlations with clinical measures evaluated whether the autism risk index was associated with autism symptom severity independent of language ability.

Results: In both samples, the autism risk index had high diagnostic accuracy (area under the curve [AUC] = 0.91 and 0.85, 95% CIs = 0.81–0.98 and 0.71–0.96), was strongly associated with Autism Diagnostic Observation Schedule–Second Edition (ADOS-2) severity scores (r = 0.58 and 0.59, $p < .001$), and not significantly correlated with language ability ($r \leq |-0.28|$, $p > .095$).

Conclusion: The autism risk index may be a useful quantitative and objective measure of risk for autism in at-risk settings. Future research in larger samples is needed to cross-validate these findings. If validated and scaled for clinical use, this measure could inform clinical judgment regarding ASD diagnosis and track symptom improvements.

Key words: autism spectrum disorder, remote eye tracking, objective measure, autism symptoms, risk

J Am Acad Child Adolesc Psychiatry 2016;55(4):301–309.

Deficits in eye gaze are a hallmark feature of autism spectrum disorder (ASD)[1,2] and are included in gold-standard diagnostic instruments.[3,4] More than a decade of research into abnormalities of eye gaze has confirmed social attention deficits as a key feature of ASD.[5-9] Across studies, diverse stimulus paradigms have elicited social attention abnormalities, ranging from decreased fixation to others' eyes[5] and social scenes[10] as early as 6 months of age, to gaze abnormalities during dyadic or joint attention bids in preschoolers[11] and older children,[12] to aberrant gaze toward dynamic social stimuli in older high-functioning individuals.[13] Subtler, but identifiable, gaze abnormalities have also been seen in family members with the broad autism phenotype.[14] This implies that eye gaze patterns, particularly those based on dynamic temporal analysis,[15] may be a promising objective risk marker of ASD as well as a quantitative measure of autism symptoms spanning the full continuum of behavior. Two recent studies provided some support for the potential discriminative value of eye gaze tracking.[16,17] In these studies, individual stimulus paradigms had modest but potentially informative discriminative value (areas under the curve [AUC] = 0.71–0.72) in separating ASD and developmental delay[17] or healthy control cases.[16] However, no published studies have evaluated whether aggregating eye tracking metrics across stimulus paradigms might show sufficient validity (AUC ≥ 0.80) to inform clinical judgment by accurately discriminating individuals with ASD from a clinically realistic comparison group.

Beyond accurate discrimination, objective measures of autism symptom severity are needed to provide quantitative assessments for tracking intervention effectiveness. At present, autism symptoms are measured using direct clinical observation, parent interview, and/or parent report.[18,19] These methods are heavily influenced by subjective perceptions, and both parent interview and clinician observation measures also require substantial training with ongoing interrater reliability checks. Parent-report questionnaires are easier to obtain and have shown validity for separating ASD and non-ASD,[20-23] but they are heavily influenced by rater biases (e.g., halo or contrast effects), measurement context, and are often conflated with other psychopathology

This article is discussed in an editorial by Dr. Frederick Shic on page 267.

Clinical guidance is available at the end of this article.

Supplemental material cited in this article is available online.

symptoms,[24,25] reducing their effectiveness in clinically challenging samples. Finally, none of the current diagnostic approaches readily produce interval-scale measurements that yield high reliability across the full range of behavior in neurotypical and ASD-affected individuals. This is true even for the Social Responsiveness Scale (SRS), where a floor effect is observed when converting low raw scores to standard scores.[26] Development of quantitative, interval-scale measures of autism symptoms, including measures of the core symptom domains of social communication/interaction (SCI) and restricted/repetitive behavior (RRB), would represent a major step forward in the technology used to capture autism symptom levels and risk for categorical ASD diagnosis.

Remote eye tracking is a promising technology for development as an objective measure of autism. In addition to literature support for gaze abnormalities in ASD, remote eye tracking is easier to calibrate and to collect in young or severely impaired children relative to traditional headgear-based eye tracking methods and other methods that require significant preparation (e.g., electroencephalogram [EEG]/event-related potential [ERP]), physical restrictions (e.g., magnetic resonance imaging [MRI]/magnetoencephalography [MEG]), and possibly sedation (e.g., MRI). Young children are familiar with watching TV, and their attention is often sufficiently captured over short intervals. When attention capture is more challenging, visual stimuli can be repeated and interspersed around short breaks, with multiple stimuli used to further enhance novelty and attention. Use of multiple stimulus paradigms also permits capture of different aspects of attention, including attention to socially appropriate targets and nonsocial/distractor targets. Relative to the massive literature examining social attention in ASD, very little research has focused on abnormalities of attention to nonsocial/distractor stimuli.[27] Inappropriate attention to nonsocial stimuli is frequently observed clinically and is an important part of the description of RRB symptoms.[28] Supporting this observation, Sasson *et al.*[27] recently demonstrated that attention to objects versus people and attention to high versus low autism-interest items were associated with greater RRB symptoms. This study suggests that it may be possible to identify characteristic gaze patterns to nonsocial stimuli that more strongly associate with RRB than SCI symptoms. Visual attention paradigms can readily integrate both social and nonsocial/distractor targets without adding time or reducing participant engagement.

The primary aim of the present research was to develop and replicate an objective measure of autism symptom level based on eye gaze tracking to social and nonsocial stimuli, hereafter referred to as the "autism risk index" (ARI). We hypothesized that ASD-affected children would show less attention to social and greater attention to nonsocial targets than children without an ASD diagnosis but with other developmental neuropsychiatric concerns (non-ASD). Based on this expectation, the ARI was created by averaging dwell times to a priori social and nonsocial target regions of interest. In initial and replication samples, the ARI was expected to show strong discrimination (AUC $\geq$ 0.80) of ASD and non-ASD cases and to be significantly related to overall autism symptom severity but not language measures.

METHOD

Samples

Participants were children 3.0 to 8.11 years of age who were referred to a tertiary care multidisciplinary ASD specialty clinic. Referrals were made by local pediatricians, following autism screening, if there was clinical concern of social deficits or ASD, or if parents or teachers had concerns. Patients were consecutively recruited at the time of the diagnostic evaluation visit (initial study: July 2014 and June 2015; replication study: August 2015 to November 2015). Gaze data were collected before the consensus diagnosis team meeting, and the research team was blinded to participant diagnosis. Procedures of this research were reviewed and approved by the Cleveland Clinic institutional review board.

Eye Tracking

Eye tracking data were collected in a quiet room adjacent to the diagnostic clinic. Data were recorded using an SMI remote eye tracker (initial study: Red-m at 120 Hz, replication study: Red250 at 60 Hz) attached to the frame of a 1280 horizontal $\times$ 1024 vertical 19-inch LCD stimulus presentation monitor. Spatial resolution of these systems was 0.1°, and average gaze position accuracies were 0.5°. The system allows for head movement (32 $\times$ 21 $\times$ 25 cm for Red-m and 32 $\times$ 21 $\times$ 30 for Red250) at a maximum distance of 75 cm. In the initial study, a 3- or 5-point calibration was obtained before the experiment. In the replication study, an initial and 4 additional 5-point calibrations were obtained at fixed times throughout the experiment (see Supplement 1, available online). Proportion net dwell time to each ROI was derived using SMI BeGaze software. Dwell time was defined as the sum of all sample durations (all fixations and saccades) falling within the ROI divided by the total stimulus time.

Visual Stimulus Battery

Stimuli were presented using SMI Experiment Center, and stimuli for the initial study were selected to represent multiple distinct types previously used in the eye gaze literature, including static facial affect, biological versus non-biological pairings, and dynamic/naturalistic scenes. Figures S1 and S2 (available online) present example stimuli created for the initial and replication studies, and Tables S1 and S2 (available online) list all stimuli and ROIs. Stimuli were presented in a single order, intermixed with attention-grabbing stimuli, gaze recalibration, and receptive language stimuli, and other stimuli not considered for the present paper. Total experiment time was approximately 7 minutes for both the initial and replication studies.

For the initial study, a priori ROIs were identified by the first author, who did not participate in data collection or diagnostic evaluations. ROIs were drawn to capture important social (faces, key body movements) and nonsocial target stimulus elements (distractors). A priori ROIs were further restricted to key time points within each stimulus based on a socially relevant action. When relevant, a priori ROIs were also designated across the total stimulus period to capture basic attention to social versus nonsocial elements. For example, in dynamic joint attention stimuli, a temporal ROI evaluated gaze to the most relevant social action (e.g., gaze-and-point to a target), but the total stimulus period was also examined to capture overall attention to the social (e.g., face) and nonsocial