Supplementary Material

Table of Contents

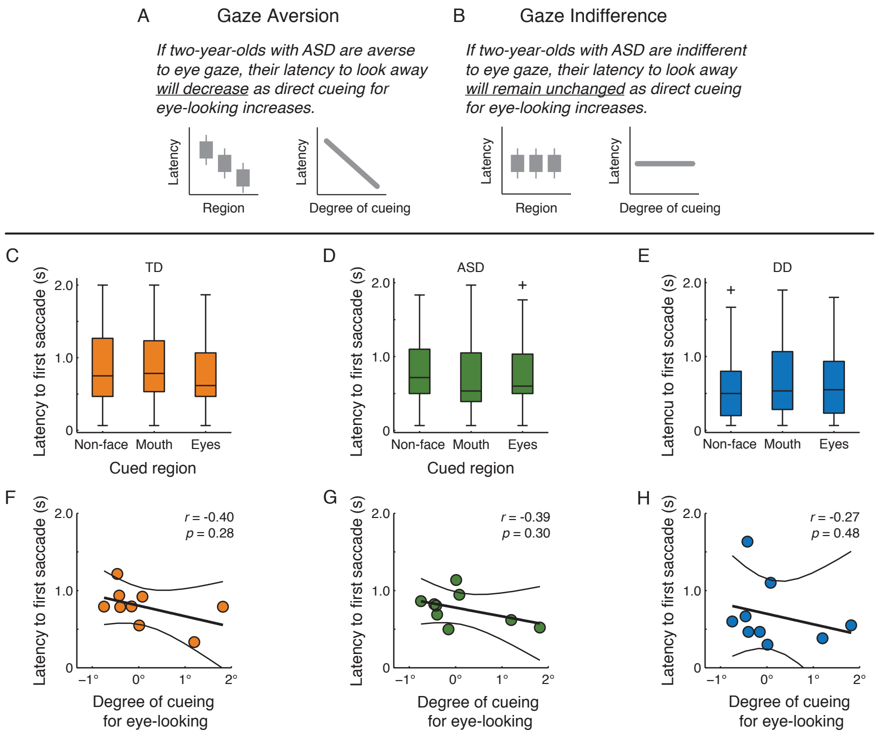

Figure S1. Following direct cueing to look at the eyes, developmentally-delayed two-year-olds do not look away more rapidly than their typically-developing peers. (A, B) Expected results for (A) the gaze aversion hypothesis and (B) gaze indifference hypothesis. (C) In typically-developing toddlers, latency to first saccade did not vary categorically as a function of the content of cued target region: eyes, mouth, or non-face. (D) In toddlers with ASD, latency to first saccade did not vary categorically as a function of the content of cued target region: eyes, mouth, or non-face. (E) In developmentally-delayed toddlers, latency to first saccade did not vary categorically as a function of the content of cued target region: eyes, mouth, or non-face. (F) In typically-developing toddlers, latency to first saccade did not vary dimensionally by degree of direct cueing for eye-looking: closer to or farther from the eyes. (G) In toddlers with ASD, latency to first saccade did not vary dimensionally by degree of direct cueing for eye-looking: closer to or farther from the eyes. (H) In developmentally-delayed toddlers, latency to first saccade did not vary dimensionally by degree of direct cueing for eye-looking: closer to or farther from the eyes. ASD = Autism Spectrum Disorder, DD = developmentally-delayed, TD = typically-developing.

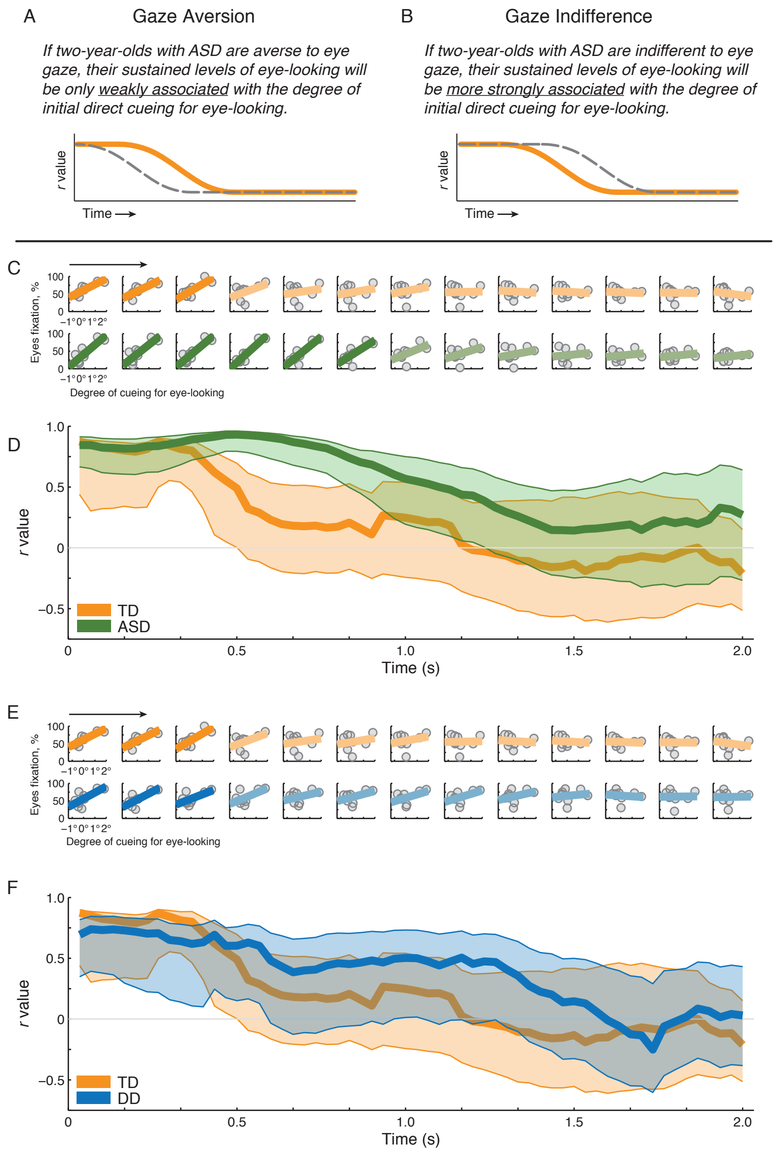

Figure S2. After direct cueing, unlike two-year-olds with ASD, developmentally-delayed two-year-olds show patterns of sustained looking comparable to typically-developing peers. (A, B) Expected results for (A) the gaze aversion hypothesis and (B) gaze indifference hypothesis. (C) Sustained effects of direct cueing are measured as the time-until-decay of a significant association between direct cueing location and percentage of looking to the cued region. For discrete time periods, plotted in 166.7 ms intervals, the association between cueing target location and percentage of fixation to cued region is plotted as scatter plots with regression lines for typically-developing toddlers (top row, orange) and for toddlers with ASD (bottom row, green). Regression lines for time periods with significant associations ($p < 0.05$) are plotted with more saturated shades for each group. (D) Measured continuously, the association (r value of correlation) between degree of direct cueing for eye-looking and percentage of eye-looking shows rapid decline in typically-developing toddlers (orange) and slower decline in toddlers with ASD (green). Shaded regions show bootstrapped 95% confidence intervals. Despite equivalent latencies to first saccade, toddlers with ASD show a stronger, more persistent association between cueing target location and subsequent fixation location, looking more at the eyes when cued for eye-looking and more at the mouth when cued for mouth-looking—an indicator of relative insensitivity to the content at the cued target location in toddlers with ASD. (E) For discrete time periods, the association between cueing target location and percentage of fixation to cued region is plotted as scatter plots with regression lines for typically-developing toddlers (top row, orange) and developmentally-delayed toddlers (bottom row, blue). Regression lines for time periods with significant associations ($p < 0.05$) are plotted with more saturated shades for each group. (F) Measured continuously, the association (r value of correlation) between degree of direct cueing for eye-looking and percentage of eye-looking shows comparable rates of decline in typically-developing toddlers (orange) and developmentally-delayed toddlers (blue). In contrast to the more persistent effect in toddlers with ASD, the duration of the effect of cueing target location on subsequent fixation patterns does not differ between typically-developing and developmentally-delayed, non-ASD toddlers. ASD = Autism Spectrum Disorder, DD = developmentally-delayed, TD = typically-developing.

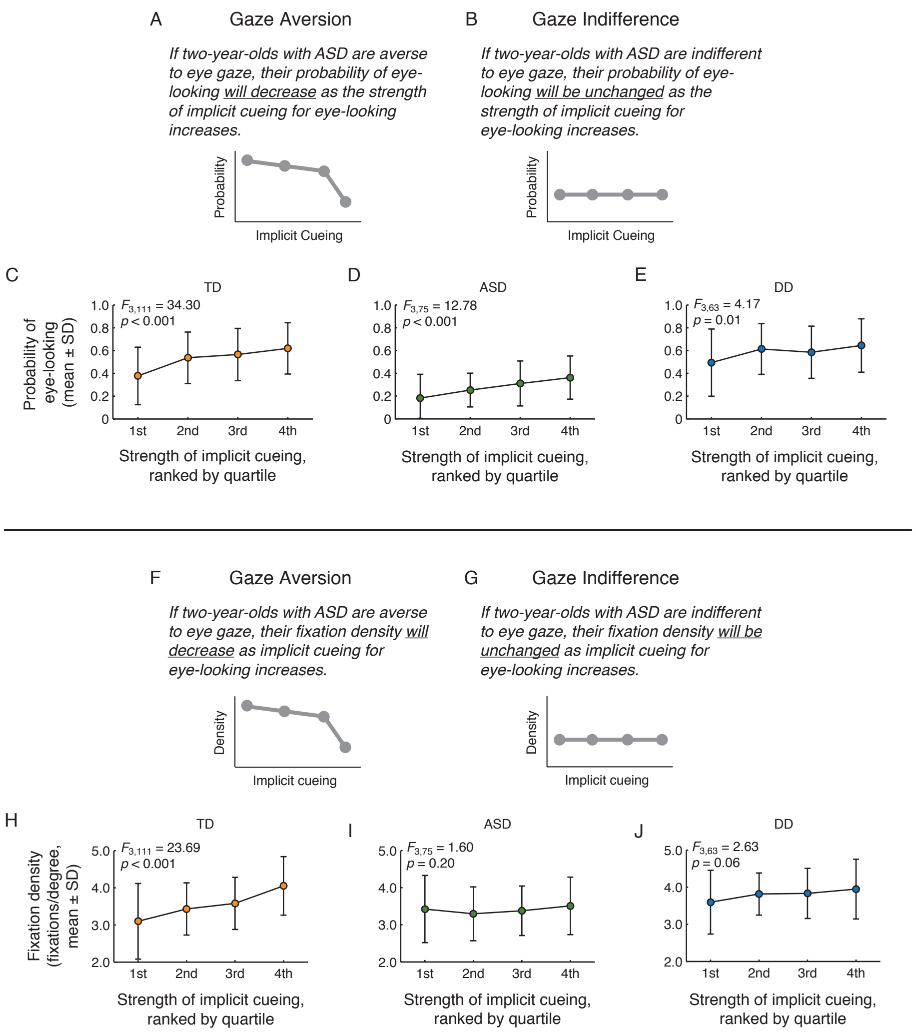

Figure S3. At moments with strongest implicit cueing for increased eye-looking, developmentally-delayed two-year-olds are more likely to look at the eyes. (A, B) Expected results for eye-looking in response to implicit cueing for eye-looking based on (A) the gaze aversion hypothesis and (B) gaze indifference hypothesis. (C) In typically-developing toddlers, implicit cueing for increased eye-looking, ranked by quartiles, positively predicts increased eye-looking among independent typically-developing children using leave-one-out cross-validation. (D) In toddlers with ASD, implicit cueing for eye-looking does not result in gaze aversion; instead, a significant increase in probability of eye-looking is observed. (E) In developmentally-delayed toddlers, increased implicit cueing for eye-looking is associated with a significant increase in probability of eye-looking. (F-J) In contast to the pattern observed in two-year-olds with ASD, fixation density in developmentally-delayed two-year-olds is related to levels of implicit cueing for increased eye-looking. (F, G) Expected results for fixation density in response to implicit cueing for eye-looking based on (F) gaze aversion hypothesis and (G) gaze indifference hypothesis. (H) In typically-developing toddlers, implicit cueing for increased eye-looking, ranked by quartiles, positively predicts greater fixation density; as implicit cueing increases, typically-developing toddlers are more likely to fixate on the same location at the same time. (I) In toddlers with ASD, fixation density is unrelated to levels of implicit cueing for eye-looking, indicating that they do not avert gaze to peripheral locations during periods of increased implicit cueing for eye-looking. (J) In developmentally-delayed toddlers, implicit cueing for increased eye-looking trends towards positively predicting greater fixation density, similar to the pattern observed in typically-developing toddlers. Error bars represent standard deviations. ASD = Autism Spectrum Disorder, DD = developmentally-delayed, TD = typically-developing.

Table S1 l Participant characterization data and percentage of visual fixation time to regions-of-interest

	ASD Group[a] (N = 26)		TD Group[a] (N = 38)		DD Group[a] (N = 22)				Pairwise *p* values		
	N		N		N		Test statistic	*p* value	ASD vs TD	ASD vs DD	TD vs DD
Sex							$X^2 = 1.36$	0.51	0.27	0.77	0.46
Male	21		26		17						
Female	5		12		5						
Diagnosis											
Autistic Disorder	19										
PDD-NOS[b]	7										
	Mean	SD	Mean	SD	Mean	SD	Test statistic	*p* value	ASD vs TD	ASD vs DD	TD vs DD
Age, months	**27.1**	6.5	**24.1**	8.0	**23.5**	7.3	$F_{2,83} = 1.68$	0.19	0.37	0.32	>0.99
Nonverbal function[c], months	**21.6**	9.0	**24.8**	9.8	**21.0**	8.0	$F_{2,83} = 1.58$	0.21	0.51	>0.99	0.37
Verbal function[d], months	**16.3**	12.4	**24.8**	9.2	**14.8**	9.0	$F_{2,83} = 8.76$	<0.001	0.005	>0.99	0.001
ADOS[e] Social Score	**9.9**	3.5			**6.0**	6.3	$t_{1,45} = 6.41$	<0.001			
Eyes, %	**28.4**	14.9	**51.5**	19.9	**51.1**	23.6	$F_{2,83} = 12.55$	<0.001	<0.001	<0.001	>0.99
Mouth, %	**42.7**	21.6	**24.7**	20.0	**25.9**	19.2	$F_{2,83} = 6.77$	0.002	0.002	0.017	>0.99
Body, %	**14.2**	8.0	**12.3**	8.1	**14.4**	10.9	$F_{2,83} = 0.52$	0.59	>0.99	>0.99	>0.99
Object, %	**14.8**	13.3	**11.4**	10.5	**8.6**	8.1	$F_{2,83} = 1.96$	0.15	0.68	0.16	>0.99

[a] ASD = Autism Spectrum Disorder, TD = typically-developing, DD = developmentally-delayed
[b] PDD-NOS = Pervasive Developmental Disorder, Not Otherwise Specified
[c] Verbal function, age-equivalence score in months on the Visual Reception subtest of the *Mullen Scales of Early Learning*
[d] Nonverbal function, age-equivalence score in months on the the Receptive and Expressive Language subtests of the *Mullen Scales of Early Learning*
[e] ADOS Social Score, total score on the social algorithm of the *Autism Diagnostic Observation Schedule* (higher scores denote higher levels of social disability)

Methods

Participants

With the written informed consent of their parents or legal guardians, eighty-six children participated in and completed the experimental protocol. The research protocol was approved as non-significant risk by the Institutional Review Board at Yale University School of Medicine. Data collection occurred in the Autism Program of the Yale Child Study Center, New Haven, CT. Families were free to withdraw from the study at any time.

The 86 participants included 26 toddlers with Autism Spectrum Disorder (ASD), 38 typically-developing toddlers, and 22 developmentally-delayed toddlers without ASD. Children with ASD and non-ASD developmental delays were consecutive referrals to a diagnostic clinic, with experimental procedures collected at the time of each child's initial diagnosis. All children had normal or corrected-to-normal vision and no history of major visual or auditory impairment (history of such impairments were exclusionary criteria for participation). All groups were matched on sex ratio, chronological age, and nonverbal cognitive ability, measured by mental age equivalents obtained from the Visual Reception subtest of the Mullen Scales of Early Learning (Mullen; 1). As expected, given hallmark language delays in ASD (2), the ASD group and typically-developing group differed on verbal functioning, measured by the average of mental age equivalents obtained through the Receptive and Expressive Language subtests of the Mullen (Table S1). To evaluate whether our findings were specific to ASD diagnostic status rather than level of verbal functioning, the ASD group and developmentally-delayed groups were matched on verbal functioning; both groups demonstrated significantly delayed verbal functioning relative to typically-developing peers (Table S1).

To qualify for inclusion in the ASD group, children met the following three inclusionary criteria: (1) they met criteria for Autistic Disorder or ASD on the *Autism Diagnostic Observation Schedule* (3), Module 1; (2) they met criteria for Autistic Disorder or ASD on the *Autism Diagnostic Interview – Revised* (4); and (3) they received a diagnosis of Autistic Disorder (21 of 28 children) or Pervasive Developmental Disorder-Not Otherwise Specified (7 of 28 children) by two experienced clinicians upon independent review of all available clinical data, including standardized testing and video of the diagnostic examination. Diagnostic guidelines followed DSM-IV-TR criteria (2), and all children would also meet criteria for ASD per current, DSM-5 criteria (5). Of the 26 toddlers with ASD included in the current study, 20 toddlers had follow-up clinical characterization data available in order to confirm diagnosis. All 20 toddlers continued to meet diagnostic criteria for ASD at follow-up. The majority of toddlers received follow-up evaluations at approximately 36 months.

To qualify for inclusion in the typically-developing group, children met the following inclusionary criteria: (1) they exhibited no developmental delays (measured as any single delay of >2SD or as 2 delays of >1.5SD in the Visual Reception, Receptive, or Expressive Language subscales of the Mullen); (2) they had no known genetic syndrome; and (3) they had no family history of ASD in first or second degree relatives.

To qualify for inclusion in the developmentally-delayed group, children met the following inclusionary criteria: (1) they exhibited significant developmental delays (measured as an single delay of >2SD or as 2 delays of >1.5SD in the Visual Reception, Reception, or Expressive Language subscales of the Mullen); (2) the possibility of an ASD diagnosis was ruled out by two experienced clinicians upon independent review of all available clinical data, including standardized testing and video of the diagnostic examination.

All aspects of the experimental protocol were performed by personnel blinded to diagnostic status of the children. All diagnostic measures were administered by trained clinicians blinded to the results of experimental procedures.

Experimental Procedures and Stimuli

Children were accompanied at all times by a parent or primary caretaker. Eye-tracking was accomplished by a video-based, dark pupil/corneal reflection technique with hardware and software created by ISCAN, Inc. (Woburn, MA, USA). Data collection cameras were remotely-mounted with data collected at a rate of 60 Hz. The system was mounted within a wall panel beneath the stimuli presentation monitor, concealed from the child's view by an infrared filter.

After entering the experimental testing room and being buckled into a car seat on a pneumatic lift, children were positioned so that viewing height and distance (approximately 76 cm) from the stimuli presentation monitor were standardized for all participants. The stimuli presentation monitor was a 20-in (50.8-cm) computer monitor (refresh rate of 60 Hz non-interlaced). Lights in the room were dimmed so that only the stimuli presentation monitor could be easily seen. Audio was played through a set of concealed speakers. The experimenter was hidden from the child's view by a curtain, but was able to observe the child at all times using a live video feed.

A five-point calibration method was used, presenting spinning and/or flashing points of light as well as cartoon animations, ranging in size from 1° to 1.5° of visual angle, on an otherwise blank screen, all with accompanying sounds. The calibration routine was followed by verification of calibration in which more calibration targets were presented at any of nine on-screen locations. Throughout the remainder of the testing session, calibration targets were shown between experimental videos to measure possible drift in accuracy. In the case that drift exceeded 3°, data collection was stopped and measures of the child's eye movements relative to fixation locations were recalibrated before further videos were presented. Across trials, calibration accuracy was not significantly different between groups ($F_{2,354} = 0.49$, $p = 0.61$).

Data Processing

Most aspects of data acquisition and all aspects of coding and data processing were automated to ensure separation between diagnostic characterization and the experimental protocol. Analysis of eye movements and coding of fixation data were performed with in-house software. Non-fixation data, comprising saccades, blinks, and off-screen fixations, were automatically identified in the first phase of analysis. Saccades were identified based on eye movement velocity, using a threshold of 30° per second. Blinks were identified as in (6). Off-screen fixations (i.e., when a child looked away from the video screen) were identified by fixation coordinates beyond the possible screen bounds. Across 148.5 s of total viewing data, measures of fixation time (as percentage of total time spent fixating on the stimuli presentation monitor) were not significantly different across typically-developing, ASD, and developmentally-delayed groups (typically-developing, 65.8% (13.9); ASD, 65.6% (13.1); developmentally-delayed, 63.6% (15.5), data given as mean (standard deviation) representing percentage of total viewing time; $F_{2,83} = 0.40$, $p = 0.67$). Measures of non-fixation data were also not significantly different across groups for saccades (typically-developing, 14.6% (5.0); ASD, 15.9% (6.1); developmentally-delayed, 15.5% (8.1); $F_{2,83} = 0.61$, $p = 0.54$), blinks (typically-developing, 4.8% (6.7); ASD, 3.9% (3.1); developmentally-delayed, 4.4% (5.1); $F_{2,83} = 0.41$, $p =$

.

0.66), or off-screen fixations (typically-developing, 14.8% (9.7); ASD, 14.7% (10.7); developmentally-delayed, 16.5% (13.1); $F_{2,83} = 0.40$, $p = 0.67$).

Eye movements identified as fixations were coded relative to four regions-of-interest (ROIs) defined within all video stimuli: eyes, mouth, body (neck, shoulders, and contours around the eyes and mouth, including hair), and object (background setting and inanimate objects). ROIs were hand-traced for all video frames (4,456 frames) and were then stored as binary bitmaps (via software written in MATLAB). Automated coding of fixation time to each ROI consisted of a numerical comparison of each child's coordinate fixation data against the bitmapped ROIs. Percentage of fixation time on each ROI was calculated relative to an individual's total fixation time. Between-group comparisons were calculated using a two-sample *t*-test.

Experiment 1: Response to Direct Cueing for Eye-Looking

In the first experiment, children were cued to fixate on a target location in each of 13 experimental trials. The cueing target was a radially symmetric, rotating circle, accompanied by a chiming sound. The cueing target size was approximately 1.5° in viewing angle, was presented for 3,100 ms, and was composed of 6 high-contrast alternating sectors colored blue or white (60° in sector angle size). Cueing target presentation was followed immediately by onset of the experimental stimulus. In all trials, in order to place equal demands on extraocular muscles (so as not to bias reaction times as a function of eccentricity of cueing location), cueing target location was standardized at the center of the stimuli presentation monitor.

Single trials in which an individual child did not look at the cueing target location (*i.e.*, cueing was unsuccessful) were excluded from analysis. Unsuccessfully cued trials were identified for each individual on each trial by 1 of 2 automated checks: measure of whether a child's fixation location was off-screen at the start of the trial, or measure of whether a child's fixation location was significantly farther from the center of the cueing target than all other children's fixations for that trial ($p < 0.05$, *i.e.*, identified as a significant outlier by the Tukey method). Across all typically-developing participants and all trials, individual data in 16 trials were unsuccessfully cued. Across all ASD participants and all trials, individual data in 14 trials were unsuccessfully cued. Across all developmentally-delayed participants and all trials, individual data in 12 trials were unsuccessfully cued. In each case when an individual's fixation location identified an unsuccessfully cued trial, data from the remainder of successfully cued individuals could be analyzed (so that all experimental trials could be analyzed, even if some individual children were not successfully cued on a given trial). In total, analyses included 313 successfully cued trials from typically-developing participants, 219 from ASD participants, and 161 from developmentally-delayed participants. Percentage of successfully or unsuccessfully cued trials was compared across groups using a Pearson chi-squared test; there was no significant difference in between-group measures of successfully vs. unsuccessfully cued trials: $X_2^2 = 1.18$, $p = 0.55$. All analyses were also repeated with unsuccessfully cued fixations *included*, with no material change in results.

The location of face stimuli varied with respect to the cueing target location in each trial. Degree of cueing was quantified both categorically (cueing for fixation on categorical eyes, mouth, or non-face regions) and also dimensionally within the face-cueing trials, quantified as the vertical distance from cueing target location to the center of the nose in degrees of visual angle (**Figure 1B**; range: -1.1° to 2.7°). Cueing targets relative to face and non-face locations were presented in pseudo-random order, with no systematic relationship between trial order and degree of cueing for eye-looking (Spearman correlation; $r_s = 0.17$, $p = 0.67$). As a further control

.

for learning effects relative to trial order, toddlers in the ASD group were randomly assigned to view trials in the original pseudo-random order ($N = 14$) while others ($N = 12$) viewed the stimuli in reverse pseudo-random order. There were no significant effects of trial order on experimental measures, and the two subsets of children were not significantly different in clinical characteristics; results were thus collapsed across forward and backward pseudo-random presentations.

The experimental stimuli were 32-bit color videos, 640x480 pixels in resolution, presented full-screen at the rate of 30 frames per second. Videos included a single (mono) channel audio track sampled at 44.1 kHz. Videos with cueing for fixation on face locations, comprising 9 of 13 trials, presented an actress looking directly into the camera and portraying the role of a caregiver, speaking to the viewing participant in toddler-directed speech (**Figure 1C**). Caregivers were filmed in front of a background that approximated a child's room, including colorful pictures and shelves of toys and stuffed animals. Videos with cueing for fixation on non-face locations, comprising 4 of 13 trials, presented children playing together in naturalistic outdoor settings.

Cueing effects on latency to first saccade. Latency to first saccade was calculated in milliseconds (ms) for each trial (**Figure 2C-D, Figure S1E**). Analyses focused on saccades occurring within the first 2.0 s of trial onset to ensure that reaction time measures were in response to the stimulus at onset rather than to later content. Because the distributions of latencies to first saccade were not normally distributed in the typically-developing, ASD, or developmentally-delayed group (tested by one-sample Kolmogorov-Smirnov test; typically-developing, $k = 0.98$, $p < 0.001$; ASD, $k = 0.98$, $p < 0.001$; developmentally-delayed, $k = 0.98$, $p < 0.001$), log-normalized values were used for between-group comparisons. Group mean latencies across all trials and across all three groups were compared using a one-way ANOVA. To compare latency to first saccade categorically across cueing locations (eyes, mouth, or non-face regions), group mean latencies were used to calculate a one-way ANOVA within each group (**Figure 2C-D, Figure S1E**). In trials in which participants were physically cued to fixate on a face, group mean latencies in each trial were also assessed dimensionally and used to calculate a Pearson correlation coefficient testing for association between latency to first saccade and degree of direct cueing for eye-looking (**Figure 2E-F, Figure S1H**).

Post-cueing effects on sustained levels of eye-looking. To assess the post-cueing effects on sustained levels of eye-looking, during and after the first saccade, we quantified level of eye-looking at each moment in time throughout the first 2.0 s following trial onset (an interval selected as double the duration of the groups' mean reaction time to first shift gaze). While the first analysis (**Figure 2, Figure S1**) tested the effects of cueing location on latency to first saccade, the second analysis (**Figure 3, Figure S2**) tested for effects of cueing location on sustained and subsequent fixations: the initial fixation location was predetermined by the location of the cueing target, but *subsequent* fixation locations and *sustained* levels of eye-looking were determined by the participants themselves, freely shifting their gaze. Level of eye-looking was calculated as duration of time spent fixating on the eye region relative to the total duration of time fixating on-screen within a centered moving window of 433 ms (window width selected as the mean fixation duration across all participants).

The effects of initial cueing target location on sustained levels of eye-looking were measured by calculating the Pearson correlation coefficient, at each moment in time, between level of eye-looking and degree of initial eyes cueing. Scatter plots are shown in **Figure 3C** (typically-developing and ASD groups) and **Figure S2E** (typically-developing and developmentally-delayed groups). Measures were calculated at each point in time during the first 2.0 s across all trials; estimates of the *r* values over time are plotted in **Figure 3D** (typically-developing and ASD groups) and **Figure S2F** (typically-developing and developmentally-delayed groups). To test for a between-group difference in the duration of the physical priming effect, we calculated bootstrapped group means and 95% confidence intervals. In each of 5,000 typically-developing, ASD, and developmentally-delayed groups randomly resampled (with replacement) from the original samples, we calculated the Pearson correlation coefficient at each moment in time. We constructed 95% confidence intervals for each group at each moment based on the 2.5^{th} and 97.5^{th} percentiles of the Pearson correlation coefficient values across all 5,000 resampled groupings. Examining the degree of overlap in the typically-developing and ASD groups' confidence intervals as well as overlap in the typically-developing and developmentally-delayed groups' confidence intervals allowed us to determine whether differences observed in the duration of the direct cueing effect passed the threshold expected by chance. We determined the duration of the direct cueing effect in each group based on the time from trial onset until the time at which the lower bound of the 95% confidence interval was no longer significantly different from 0 (two-sided α level of 0.05).

Experiment 2: Response to Implicit Cueing for Eye-Looking

In the second experiment, children were presented with 9 videos that contained implicit, time-varying cues for eye-looking (**Figure 1C**). As with Experiment 1, video trials were presented in pseudo-random order. Each video presented an actress looking directly into the camera and portraying the role of a caregiver, speaking to the viewing participant in toddler-directed speech.

Implicit cueing for eye-looking. We identified implicit cueing for increased eye-looking by analysis of normative data: Given previous research in typically-developing children demonstrating intrinsic engagement with and responsiveness to underlying social cues from an actress' eyes (7,8), we measured the probability of typically-developing 2-year-olds looking at the eyes while freely viewing the video stimuli in 33.3 ms intervals throughout each video (e.g., 80% of typically-developing 2-year-old participants were looking at the eyes at time t = 10.000 s, but only 60% of were looking at the eyes at time t = 12.033 s, yielding probabilities of 0.8 and 0.6, respectively). Probability of eye-looking varied between 0 and 1 across video trials. The strength of implicit cueing for eye-looking was indexed as quartiles defined across all videos (**Figure 1D**). Because the start of each video was preceded by direct cueing, measures of implicit cueing specifically excluded the first 1.5 s of each video (during which time, per **Figure 3**, the effects of physical cueing persist).

To test the validity of these implicit cues for eye-looking, we used leave-one-out cross validation (LOOCV; 9). Rather than define implicit cueing for eye-looking on the basis of *all* typically-developing participants' data (N = 38), implicit cueing for eye-looking was defined on the basis of all typically-developing participants' data *excluding the data of one (left-out) participant* (N = 37); the left-out participant's data was then used as an independent test case for testing the timeline of implicit cueing defined by the *other* participants' data (i.e., if other participants look more at the eyes at time t_1 and less at the eyes at time t_2, is an independent participant also likely to look more at the eyes at time t_1 and less at the eyes at time t_2?). This procedure was repeated for each individual typically-developing participant, so that each typically-developing participant's data was independently analyzed and used to test the predictive validity of implicit cues defined by other participants' increased eye-looking. To test the validity of this measure of implicit cueing for eye-looking, we then computed within-group repeated measures ANOVAs across all LOOCV comparisons to measure whether the timing of eye-looking for each typically-developing child was successfully predicted by the other typically-developing children's looking. The results showed high predictive validity across all cross-validations ($F_{3,111} = 34.30$, $p < 0.001$).

Implicit cueing effects on probability of eye-looking. **Figure 4C** shows a significant effect of the strength of implicit cueing for eye-looking on the actual probability of eye-looking of independent typically-developing toddlers. To assess the effect of implicit cueing in ASD toddlers, the probability of spontaneous eye-looking of ASD and developmentally-delayed toddlers was tested with a within-group repeated measures ANOVA relative to the strength of implicit cueing for eye-looking, ranked by quartile, determined using the full sample of typically-developing toddlers (**Figure 4D, Figure S4E**).

Implicit cueing effects on fixation density. To quantify the spatial distribution of fixations in response to implicit social cueing, we measured fixation density by kernel density analysis (10) at each moment in time throughout the duration of all video trials. In this method, viewers' fixation locations form a density function that quantifies the distribution of onscreen fixation locations. The fixation location of each individual viewer, relative to the other viewers' fixation locations, is thereby associated with a density value. This method enabled us to measure whether fixation locations were narrowly distributed near the same location (e.g., the eyes) or widely distributed across different locations at each moment in time and specifically during moments of implicit social cueing for increased eye-looking. The effect of variation in implicit cueing for eye-looking on fixation density was assessed in typically-developing, ASD, and developmentally-delayed toddlers using within-group repeated measures ANOVA (**Figure 4G-H, Figure S4J**).

RESULTS

Developmentally-Delayed Group

Throughout free-viewing of the video scenes, developmentally-delayed toddlers exhibited an overall level of eye-looking comparable to typically-developing peers (**Table S1**; $p > 0.99$) and significantly higher than ASD peers ($p < 0.001$). The result replicated past findings (11) and indicated that the developmentally-delayed group did not show diminished attention to others' eyes and, consequently, did not meet the basic behavioral criterion for either gaze aversion or indifference.

In Experiment 1 assessing response to direct cueing for eye-looking, the developmentally-delayed group showed no significant relationship between latency to first saccade and degree of eyes cueing measured either categorically (**Figure S1E**; $F_{2,108} = 0.76$, $p = 0.47$) or continuously (**Figure S1H**; $r = -0.27$, $p = 0.48$), a pattern comparable to that observed in the typically-developing group (**Figure S1C, S1F**). The effect of direct cueing on sustained levels of eye-looking was also comparable between the developmentally-delayed and typically-developing groups; bootstrapped confidence intervals for each group were overlapping throughout the first 2,000 ms of the video scenes (**Figure S2F**). The significant association between eyes cueing and level of eye-looking persisted in the developmentally-delayed group for 567 ms, similar to the 500 ms effect observed in the typically-developing group and about 2 times less than the 1,233 ms effect observed in the ASD group (**Figure S2D**).

In Experiment 2 assessing response to implicit cueing for eye-looking, the developmentally-delayed group showed a significant main effect of level of implicit cueing on the amount of eye-looking (**Figure S3E**; $F_{3,63} = 4.17$, $p = 0.01$), a pattern comparable to typically-developing peers. There was no interaction effect when compared to the typically-developing group ($F_{3,174}$ = 1.88, p = 0.14). The developmentally-delayed group also showed a trending main effect of level of implicit cueing on fixation density (**Figure S3J**; $F_{3,63}$ = 2.63, p = 0.06), though there was a significant interaction effect when compared to the typically-developing group ($F_{3,174}$ = 8.81, p = 0.004).

Across nearly all measures, including overall level of eye-looking as well as response to both direct and implicit cueing for eye-looking, the developmentally-delayed group's gaze patterns were comparable to those of the typically-developing group, providing additional evidence that differences observed in the ASD group were due to diagnostic status rather than to verbal developmental level.

.

EXTENDED DISCUSSION

Developmental Learning Model of Later-Life Gaze Avoidance in ASD

Although our results contradict the hypothesis that children with ASD actively avoid looking at the eyes in early life, and thus refute the theory of gaze aversion as a causal mechanism of social disability in ASD, we do not suggest that all other evidence supportive of gaze aversion is necessarily invalid. As described in the Discussion section of the main text, existing accounts of gaze-related anxiety are generally from later childhood or adulthood.

One possible explanation for cases of developmental progression *away* from passive omission of eye-looking, towards active omission, is that some children with ASD may learn, over time, to avoid eye contact through associative learning: associating a stimulus with an event as in classical conditioning, a type of learning that has often been noted as a relative strength of children with ASD (12). The stimulus in this case would be a person approaching and making eye contact to engage the child in a particular activity. If that activity is non-preferred (e.g., eating a particular food, performing a less enjoyable daily living skill), then the learning environment is primed for the child to associate the approach of another person making eye contact with the non-preferred activity and with the cessation of other, preferred activities. In their attempts to engage a child, most people deploy behaviors that are *normatively* beneficial for engaging children, one of which is more exaggerated bids for eye contact (13). For a child with ASD, presenting additional bids for eye contact is not likely to become more engaging or more meaningful, but may instead increase the degree to which eye contact is paired with the non-preferred activity. As a result, the child's behavioral response strategy—gaze omission—is more likely to assume habitual qualities and can then, as a conditioned response, be elicited by others' attempts at social engagement (14).

The implication of this model is that some older individuals with ASD may learn to avoid eye gaze—not by virtue of hyperarousal or understanding of gaze cues—but by way of cue-based associative conditioning in which another's eye gaze is repeatedly paired with a non-valued outcome, ultimately eliciting reflex-like omission of eye-looking (14). This type of learning is distinct from Skinnerian action-outcome learning in which an action is performed in order to produce a known (expected) outcome. Typically-developing children are continuously engaging in forms of behavior acquired through action-outcome learning that are highly social, responding to and engaging with other individuals in order to achieve a reinforcing outcome (e.g., social attention, smiles, physical contact), and modifying their social engagement strategies based on a history of behaviors previously resulting in positive social interaction (15). For children with ASD, existing evidence suggests that the "outcome" in this model of action-outcome learning does not have the same reward value that it has for TD children (16). Diminished reward value of an outcome alone is a means of disrupting action-outcome learning. Moreover, if the relationship between action and outcome is not understood, this type of learning cannot happen (17). Instead, stimulus-based associative learning mechanisms stand in as a framework by which to develop and engrain behavioral response strategies, a phenomenon demonstrated extensively in the animal literature (17).

Limitations

There are several limitations to the current study. Individual differences in social sensitivity to others' eyes among children with ASD are both possible and likely, particularly as children age and gain variable experience in social interactions. Although our results indicate that gaze indifference rather than gaze aversion is the most prevalent underlying mechanism of atypical eye contact for a majority of children with ASD at time of initial diagnosis, it is possible that our findings may not hold true across the entire spectrum in ASD, such as in cases of non-idiopathic ASD (e.g., ASD with comorbid Fragile X Syndrome; 18). It is also possible that older adolescents and adults with ASD may omit and avoid eye contact as a learned response via the stimulus-based associative learning mechanism described above. A longitudinal study evaluating attention to others' eyes from time of initial diagnosis to adolescence and adulthood will be critical to evaluate this learning theory-based model of the development of gaze omission in ASD.

References

1. Mullen EM. Mullen Scales of Early Learning: AGS Edition. Circle Pines, MN: American Guidance Service; 1995.
2. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th editio. Washington, DC: American Psychiatric Association; 2004.
3. Lord C, Rutter M, DiLavore PC, Risi S. Autism Diagnostic Observation Schedule. Los Angeles, CA: Western Psychological Services; 2002.
4. Rutter M, LeCouter A, Lord C. Autism Diagnostic Interview-Revised. Los Angeles, CA: Western Psychological Services; 2003.
5. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th editio. Washington, DC: American Psychiatric Association; 2013.
6. Shultz S, Klin A, Jones W. Inhibition of eye blinking reveals subjective perceptions of stimulus salience. Proc Natl Acad Sci. 2011;108(52):21270–5.
7. Senju A, Tojo Y, Dairoku H, Hasegawa T. Reflexive orienting in response to eye gaze and an arrow in children with and without autism. J Child Psychol Psychiatry. 2004 Mar;45(3):445–58.
8. Chawarska K, Klin A, Volkmar F. Automatic Attention Cueing Through Eye Movement in 2-Year-Old Children With Autism. Child Dev. 2003;74(4):1108–22.
9. Stone M. Cross-validatory choice and assessment of statistical predictions. J R Stat Soc Ser B. 1974;36(2):111–47.
10. Silverman BW. Density Estimation for Statistics and Data Analysis. London: Chapman and Hall; 1986.
11. Jones W, Carr K, Klin A. Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. Arch Gen Psychiatry. 2008 Aug;65(8):946–54.
12. Boucher J, Mayes A, Bigham S. Memory in Autistic Spectrum Disorder. 2012;138(3):458–96.
13. Kleinke CL. Gaze and eye contact: a research review. Psychol Bull. 1986 Jul;100(1):78–100.
14. Yin HH, Knowlton BJ. The role of the basal ganglia in habit formation. Nat Rev Neurosci. 2006;7(6):464–76.
15. Tarabulsy GM, Tessier R, Kappas a. Contingency detection and the contingent organization of behavior in interactions: implications for socioemotional development in infancy. Psychol Bull. 1996;120(1):25–41.
16. Dichter GS, Richey JA, Rittenberg AM, Sabatino A, Bodfish JW. Reward Circuitry Function in Autism During Face Anticipation and Outcomes. J Autism Dev Disord. 2012;42:147–60.
17. Balleine BW, O'Doherty JP. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action.

.

Neuropsychopharmacology. 2010;35:48–69.

18. Cohen IL, Vietze PM, Sudhalter V, Jenkins EC, Brown WT. Parent-child dyadic gaze patterns in fragile X males and in non-fragile X males with autistic disorder. J Child Psychol Psychiatry. 1989 Nov;30(6):845–56.

.