Published in final edited form as:

J Am Acad Child Adolesc Psychiatry. 2017 July ; 56(7): 546–555. doi:10.1016/j.jaac.2017.05.005.

A Meta-Analysis of Gaze Differences to Social and Nonsocial Information Between Individuals With and Without Autism

Dr. Thomas W. Frazier, PhD,

Center for Autism, Cleveland Clinic, Cleveland, OH

Dr. Mark Strauss, PhD,

University of Pittsburgh, Pittsburgh, PA

Mr. Eric W. Klingemier, BA,

Center for Autism, Cleveland Clinic, Cleveland, OH

Ms. Emily E. Zetzer, MA,

Center for Autism, Cleveland Clinic, Cleveland, OH

Dr. Antonio Y. Hardan, MD,

Stanford University, Stanford, CA

Dr. Charis Eng, MD, PhD, and

Genomic Medicine Institute, Cleveland Clinic

Dr. Eric A. Youngstrom, PhD

University of North Carolina at Chapel Hill

Abstract

Objective—Numerous studies have identified abnormal gaze in individuals with autism. Yet only a limited number of findings have been replicated, the magnitude of effects is unclear, and the pattern of gaze differences across stimuli remains poorly understood. To address these gaps, we conducted a comprehensive meta-analysis of autism eye tracking studies.

Method—PubMed and manual search of 1,132 publications were used to identify studies comparing looking behavior to social and/or nonsocial stimuli between individuals with autism and controls. Sample characteristics, eye tracking methods, stimulus features, and regions-of-interest (ROI) were coded for each comparison within each study. Multivariate mixed-effects meta-regression analyses examined the impact of study methodology, stimulus features, and ROI on effect sizes derived from comparisons using gaze fixation metrics.

Correspondence to Thomas W. Frazier, PhD, Center for Autism, Cleveland Clinic, 2801 Martin Luther King Jr. Drive CRS10, Cleveland, OH 44104; fraziet2@ccf.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Supplemental material cited in this article is available online.

Disclosure: Drs. Strauss and Eng and Ms. Zetzer report no biomedical financial interests or potential conflicts of interest.

Results—The search revealed 122 independent studies with 1,155 comparisons. Estimated effect sizes tended to be small-to-medium, but varied substantially across stimuli and ROI. Overall, nonsocial ROIs yielded larger effect sizes than social ROIs; however, eye and whole face regions from stimuli with human interaction produced the largest effects (Hedge's $g=.47$ and $.50$, respectively). Studies with weaker study designs/reporting yielded larger effects, but key effects remained significant and medium-sized, even for high-rigor designs.

Conclusion—Individuals with autism show a reliable pattern of gaze abnormalities that suggests a basic problem with selecting socially-relevant versus irrelevant information for attention and that is persistent across age and worsens during perception of human interactions. Aggregation of gaze abnormalities across stimuli and ROI could yield clinically useful risk assessment and quantitative, objective outcome measures.

Keywords

autism spectrum disorder; eye tracking; meta-analysis; meta-regression; social information processing

INTRODUCTION

Autism spectrum disorder (ASD) is a complex neurodevelopmental disorder characterized by social communication and interaction impairments, along with restricted patterns of interest and/or repetitive behaviors.¹ Early descriptions of ASD emphasized abnormal gaze (looking at irrelevant stimuli or unusual eye contact patterns),² and gold-standard diagnostic instruments include this characteristic.^{3,4} Over the last 14 years, a large number of eye tracking studies, employing a variety of methodological characteristics, have supported deficits in attention to social information as a key feature of ASD.^{5–8} Across studies, diverse stimuli and regions-of-interest (ROIs) within stimuli have elicited social attention differences, ranging from decreased fixation to others' eyes⁵ and social scenes⁹ as early as 6 months of age, to aberrant gaze toward dynamic social stimuli in older cognitively-able individuals.¹⁰ The massive methodological and sampling heterogeneity across studies makes definitive conclusions regarding the exact nature of gaze abnormalities unclear. Recent meta-analyses have suggested small-to-medium reductions in looking to socially relevant regions, particularly eye and whole face regions,⁷ and increased gaze to less relevant regions of the stimulus (e.g., nonsocial regions),^{11,12} including extraneous objects and non-core face regions (e.g., hair, ears, etc.). However, these reviews had limited focus, emphasized single effects within studies when most studies include multiple comparisons across different stimuli and ROIs, and/or examined only one or two types of stimuli. The present meta-analysis surveyed the literature more comprehensively, coded studies for the full range of methodological characteristics, and used advanced multivariate meta-regression procedures to determine specific factors influencing gaze abnormalities in autism. The multivariate approach avoided splitting types of stimuli or ROIs into separate analyses with lower power, provided direct comparisons across stimuli and ROIs, evaluated combinations of stimulus and ROI features, and used studies reporting multiple effects as their own control.

Eye tracking studies provide a unique window into the moment-by-moment selection of information for attention and consequently for cognitive processing. As a result, quantitative

review of the eye tracking literature has the potential to inform our understanding of cognitive processing deficits observed in individuals with ASD.¹³ Several cognitive theories of ASD have been postulated, ranging from abnormalities of basic perceptual processing^{14,15} to models emphasizing higher-order social cognitive processes, such as deficits in social motivation¹⁶ or theory of mind.¹⁷ Cognitive theories can be roughly divided into those that emphasize aberrant attention to social and affective information^{16–19} (e.g., intentional attunement, theory of mind, emotional salience landscape theory), hereafter referred to as social processing theories, versus those that highlight deficits independent of social or affective information^{20,21} (e.g., executive dysfunction, global vs. local processing, multi-sensory processing)—hereafter referred to as general or nonsocial processing theories. Note that these distinctions are not absolute. For example, executive functioning and multisensory processing may be influenced by the social or affective nature of stimuli, while attention to social information can be influenced by basic sensory processes and stimulus complexity. Thus, quantitative review of eye tracking studies is not likely to provide direct evidence confirming or excluding any specific cognitive theory of autism, but it can inform our understanding of core features that are prominent in a majority of ASD cases.

The primary aim of the present meta-analysis was to estimate the magnitude and significance of gaze differences to socially relevant and nonsocial regions-of-interest in individuals with versus without ASD. Based on recent findings across a wide range of stimuli and paradigms,²² we hypothesized that gaze differences would be significant and medium-sized, with larger effect sizes seen to face regions and nonsocial regions embedded within a social context. The secondary aim was to evaluate whether gaze differences were stronger to stimuli that included social interaction and/or increased complexity (e.g. dynamic stimuli with audiovisual elements and paradigms with explicit directives). This aim could inform whether the most prominent cognitive deficits in ASD are restricted to processing of social information, whether more general cognitive deficits would be observed, or whether both types of impairments would be reflected in gaze differences. Based on the notion that ASD results from both social processing and general cognitive impairments, we hypothesized that stimuli depicting social interactions and those with increasingly complex stimulus features would result in larger gaze differences.

METHOD

Search Strategy and Inclusion Criteria

A comprehensive literature search queried the PubMed database with the final search updated January 4, 2016. Search terms included: “autism” OR “autistic” OR “pervasive developmental” AND “eye” OR “gaze” OR “fixation” OR “looking.” Manual searches perused the reference sections of retrieved studies and review articles. Studies (or effects within studies) were included if a) compared gaze between ASD and controls; b) used one of the following gaze metrics: fixation count, fixation count percent, fixation duration, fixation duration percent, or time-to-first-fixation (looking/duration lumped with fixation metrics); c) the comparison focused on a clearly-delineated social or nonsocial region-of-interest, d) the effect size was directly computable or could be readily estimated from available figures, AND e) the study included unique data not previously reported. Studies (or effects within

studies) were excluded if: a) they were based on other gaze metrics (saccades, blinks, etc.), b) data were presented in another paper in a more complete form, c) an effect size for the group comparison could not be computed from the available information, d) the study was not in English or could not be retrieved, e) gaze measurements relied on visual coding rather than direct measurement with a tracking system, f) the study focused on gaze differences to language or reading stimuli rather than social or nonsocial stimuli, or g) the study included at-risk or sibling groups instead of individuals diagnosed with ASD. Our initial search was limited to 1999 and later, but the first studies meeting inclusion criteria were identified in 2002. The final search generated 1,123 hits, with 9 additional papers identified via reference lists and recent reviews. A content expert examined each title and abstract for inclusion. Any abstracts describing an eye tracking study of autism were downloaded and the full paper reviewed. A total of 120 papers were identified from PubMed meeting inclusion/exclusion criteria. Adding the 9 manually identified papers yielded 129 papers. Three papers reporting non-unique data were excluded, while 5 papers reported unique effect sizes from samples that overlapped a primary source. Data from these 5 papers were lumped with data from the 4 primary sources. One publication reported two independent studies (Figure 1). Supplement 1, available online, presents all study references, and Table S1, available online, presents compliance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting standards for meta-analyses.²³

Study Coding

The first author (T.F.) extracted and coded each study. Sample characteristics included type of comparison group—healthy controls or developmental disability (DD), group sample sizes, average age and sex composition (%) of the autism and control groups, diagnostic system used, whether the Autism Diagnostic Observation Schedule (ADOS) or Autism Diagnostic Interview-Revised (ADI-R) was administered, and average IQ of the autism and control groups. Each study's sampling rate and points of regard used in gaze calibration were recorded and converted to ordinal codes (see Supplement 2, available online). Stimulus characteristics included: audio-visual versus visual-only (Audio-Visual), dynamic versus static (Dynamic), face only versus other more complex scenes (Face Only), human present versus other (Human), interactions between human or non-human characters (e.g., cartoon characters or animals playing) versus no interaction (Interactive), number of stimuli included in the group comparison (Number of Stimuli), and whether a directive was given to the participant (Directive) (See Supplement 3, available online). Group comparisons were based on regions-of-interest (ROIs) within stimuli. ROIs were coded as 0=nonsocial and 1=social. Social ROIs were defined as those where the participant would be expected to focus attention on that region, while nonsocial ROIs were regions without obvious socially relevant content where minimal gaze would be expected. Social ROIs were further defined into body part regions (eyes, whole face, upper face, lower face, face ratio [upper versus lower face], mouth, nose, non-body regions). Non-body regions were ROIs where the gaze expectation is to an object, such as in joint attention paradigms. Nonsocial ROIs included other face regions and other body regions. (See Supplement 4, available online, for examples of social and nonsocial ROIs). Standardized mean difference effect sizes (Hedge's g – a correction of Cohen's d for small sample bias) were either recorded directly from the study or were calculated using appropriate formulae.²⁴ For social ROIs, comparisons were coded

so that positive effect sizes indicated less looking (lower fixation count or duration) or greater time-to-first-fixation in the autism group. For nonsocial ROIs, comparisons were coded so that positive effect sizes indicated more looking (greater fixation count or duration) or lower time-to-first-fixation in the autism group.

Design and Reporting Quality Index—To evaluate whether study reporting and quality influenced effect magnitude, we computed a design and reporting quality index²⁵ (see Supplement 5, available online). Reporting quality was included as a surrogate for poor methodology beyond obvious design features. The index summed 12 yes/no variables coded from available study characteristics and included whether: groups were closely matched on age (defined as mean differences ≤ 0.5 years for children less than 5, ≤ 1 year for children 6–17, and ≤ 2 years for adults), sex was reported for the autism and control groups, groups were closely matched on sex ratio (defined as $\leq 10\%$ mean difference), the diagnostic system was reported, the ADOS and ADI-R were administered, IQ was reported for the autism and control groups, IQ was closely matched (defined as <7 standard score point mean difference), and the sampling rate and points of calibration procedure were reported. Quality and reporting features were also separated into sub-indices, and post hoc analyses examined whether these aspects differentially influenced effect magnitude.

Rater reliability—To evaluate whether study characteristics were coded accurately, a second rater (E.Z.), blinded to study hypotheses, coded study methodology, stimulus characteristics, and regions-of-interest for 10 studies representing 62 group comparisons. Interrater reliability was excellent for each variable ($K=.81-1.00$, absolute agreement=91%-100%). The disagreements occurred for the interactive versus other, dynamic versus static, and social versus nonsocial ROI codes. These disagreements were miss-codes by the blinded rater and precipitated a review of these variables across all studies. No additional changes were made to the original codes.

Statistical Analyses

Our summary effect size was Hedges' g . Standard formulas converted means, variances, and test statistics into g . Study variance estimate calculations followed standard methods.²⁶ All estimates used inverse variance weighting, and we report 95% CIs for the weighted effect sizes. Most studies reported multiple relevant effect sizes. Effect size nesting was most commonly due to reporting multiple ROIs but also occurred when different control groups were compared to the autism group for the same ROI. The *metafor* package²⁷ in R was the platform for all analyses, as it handles nested effect sizes within the same sample, permitting evaluation of key moderator variables.

Analyses used multivariate mixed effects meta-regression models with maximum likelihood estimation.²⁶ Prior to testing our aims, we first computed a baseline random effects multivariate model without predictors to determine overall heterogeneity of effect sizes and to estimate the intra-class correlation representing the strength of relationships among within-study effect sizes. Next, a methodological meta-regression model was estimated with study and sample characteristics as moderators. This model was used to determine whether any methodological features may influence effect magnitude and, therefore, should be

included in models testing hypotheses. Publication year (mean-centered at 2011), age of the autism group, control type (0=Healthy controls, 1=DD controls), design and reporting quality index (mean-centered at 8.63), and gaze measure type (fixation count as the comparator) were predictors. Because several methodological variables were not able to be coded for every study, we also computed individual meta-regressions for each of the following factors, using all predictors from the baseline model as covariates: sex and IQ of the autism group, autism-control IQ differences, sampling rate, points of calibration, and number of stimuli per comparison.

To evaluate the primary hypotheses, a conceptual meta-regression model used all stimulus factors, social versus nonsocial ROI, and the body part ROIs. Main effects for all factors and interactions (2- and 3-way) between dynamic, human, and interactive stimulus types tested whether combinations of stimulus features influenced effect magnitude. Based on the results of the methodological meta-regression, publication year, method quality, and control type were included as predictors to adjust for the influence of these characteristics and provide estimates of effect magnitude that were less biased. For this analysis, the “Other Face” ROI was chosen as the body part comparison because preliminary analyses indicated that this region had an average effect size that fell at the middle of the distribution for the body part ROIs. Model-derived predictions were generated for different stimulus and ROI types to demonstrate which factors and combinations of factor produced the largest effect sizes.

Meta-Analysis Reporting Standards²⁸ suggest estimating power when conducting meta-analyses. We estimated the lower bound of power using 122 independent samples ($\alpha=.05$, two-tailed). Power to detect small effect sizes ($g=.20$) was excellent ($1-\beta=.94$), even when large between-study heterogeneity was assumed ($\tau^2=.30$)²⁹. Power to detect significance ($\alpha=.05$, two-tailed) of individual regression coefficients was adequate (.76) if large relationships with effect magnitude ($\beta=.40$) and modest sampling variability ($SE=.15$) were present.

Sensitivity analyses examined whether parameter estimates from meta-regression analyses were well estimated, any outliers were present, and any publication bias might be operating. To evaluate whether parameters were well estimated, profile likelihood plots were created of the within-study and between-study variance estimates. To identify outliers, results of the conceptual meta-regression model were used to compute standardized residuals. The plot of standardized residuals versus observed effect sizes evaluated whether any observed effect sizes were substantially larger than predicted. Once identified, outliers were removed and the meta-regression model re-computed to determine the consistency of findings. The model was also re-estimated after adding points of calibration, which was a significant independent methodological factor. Publication bias was evaluated using the rank correlation test for funnel plot asymmetry.²⁶

RESULTS

Study Accounting

Figure 1 presents the flow diagram of the search process. We identified 122 independent studies, published in 2002–2015, containing 1,155 effect sizes (range 1–46 per study; 866

social, 289 nonsocial ROIs). Studies comprised a total of 5,033 participants (2,199 with ASD, 2,418 healthy controls, and 416 DD controls) ages 4 months to 40 years old. Tables S2–S4, available online, provide additional participant, moderator, and study information. The sample of studies includes a broad range of study-level characteristics, including IQ levels. However, studies with older, cognitively able (IQ ≥ 80) participants predominated (autism age IQ $r=.67$). Heterogeneity of effect sizes was highly significant ($Q[1154]=2949.0$, $p<.0001$), but moderate in size ($\tau=.441$, $\tau^2=.195$, between-study $\sigma^2=.098$, within-study $\sigma^2=.097$), while the intra-class correlation was substantial ($ICC=.503$), justifying multivariate modeling of within-study nesting.

Influence of Methodological Characteristics

The method quality index showed substantial variability across studies (range 1–11) with average quality being 8.63 out of 12. Meta-regression of methodological factors identified significant influence of methodological factors on effect sizes (moderator $Q[8]=33.18$, $p<.0001$). Higher method quality index score and publication year were significant independent predictors of lower effect magnitude (Table 1). Low method quality was associated with very large effects ($g\sim 1.0$), while high method quality yielded medium effects ($g\sim 0.50$) after adjusting for covariates (Figure 2A). Early publications generated large effects ($g\sim 0.80$), while recent publications produced medium-to-large effects ($g\sim 0.60$) (Figure 2B). The use of DD controls also trended toward lower effect magnitude (Figure 2C). Individual variable methodological meta-regressions for the remaining study characteristics identified that a larger number of calibration points was significantly associated with larger effect sizes. A non-significant tendency was noted for more stimuli used in the comparison to be associated with lower effect magnitude. Importantly, no other methodological characteristic had a significant relationship with effect magnitude, and effect sizes were remarkably consistent across the ages (see Figure S1A, available online) and IQ levels (Figure S1B, available online) sampled.

Influence of Stimulus and ROI Features

Stimuli depicting human interaction produced larger effect sizes than other stimuli (Table 2; Figure 3A). Gaze differences to nonsocial ROIs were stronger overall than gaze differences to social ROIs ($g=.11$, $SE=.05$), and this difference remained consistent across age. However, there was substantial variability in effect magnitude across social ROIs, with eye and whole face regions being associated with the largest effect sizes (Figure 3B). No other stimulus characteristic or ROI feature moderated effect magnitude. Contrary to prediction, more complex (Face Only vs. Other $p=.866$), dynamic (Dynamic vs. Static $p=.159$), audio-visual (Audio-Visual vs. Visual-Only $p=.810$) stimuli with explicit performance directives (Directive vs. No Directive $p=.684$) did not produce larger effects.

Sensitivity Analyses

Meta-regression variance parameters were well estimated (see Figure S2, available online). Several outlier effect sizes were detected ($k=22$). These effect sizes were substantially larger (residual $>2SD$) than predicted by meta-regression (see Figure S3A, available online). The rank test for publication bias confirmed this pattern (Kendall's $\tau=.176$, $p<.0001$). Inspection of these effects indicated that they were more likely to be from early studies ($M=2009$ vs.

2011; $t[22]=2.37$, $p=.027$) with lower method quality scores ($M=6.45$ vs. 8.67 ; $t[22]=5.44$, $p<.0001$), including half with method quality 6. Removing these outliers and re-running the model did not alter the pattern of substantive findings. Similarly, adding points of calibration as an additional methodological factor to the model did not alter results (see Table S5, available online). After removing outliers and adjusting for points of calibration, publication bias remained significant but was substantially reduced (see Figure S3B, available online; Kendall's $\tau=.089$, $p=.0002$).

Given that methodological factors influenced effect magnitude and that outliers were from lower quality studies and earlier publication years, effect magnitude was predicted for studies with low rigor (defined as Year=2005 and Method Quality Index<5), moderate rigor (defined as Year=2011 and Method Quality Index=8.63), and high rigor (defined as Year=2013 and Method Quality Index=10). Figure 2D shows a major effect of study rigor with more than a doubling of effect magnitude in low rigor studies relative to high rigor studies. However, it is important to note, that there is less distinction between moderate and high rigor studies and even high rigor studies had a non-trivial and statistically significant predicted effect size. Additionally, when methodological rigor was high, eye and whole face ROIs within stimuli depicting human interaction produced highly significant, medium effect sizes (Eye $g=0.47$, $SE=0.12$; Whole Face $g=.50$, $SE=0.12$), with nonsocial ROIs only slightly lower ($g=.41$, $SE=0.10$).

DISCUSSION

To our knowledge, this is the only quantitative review of gaze differences between autism and control participants to comprehensively and simultaneously evaluate all stimulus types. Gaze abnormalities were present across a wide array of stimuli and ROIs, but the largest effects were observed for stimuli depicting human interaction and for ROIs that are crucial for appropriate perception of social and emotional context. Specifically, individuals with autism showed greater attention to nonsocial regions that lack important social cues and less attention to eye and whole face regions embedded within human interaction stimuli that are crucial to accurate social perception. This pattern confirms the widely-held notion that social processing differences are a key feature across the full spectrum of autism. Intriguingly, more complex, dynamic, audio-visual stimuli and paradigms that involved explicit performance directives did not produce larger effects. This suggests that adding perceptual complexity to stimuli or increasing cognitive load during attention does not increase gaze abnormalities in autism. While deficits in multisensory processing, executive dysfunction, and global-local processing biases are certainly present in many individuals with autism,^{20,21,30} the present observations suggest that general (non-social) cognitive deficits may be less core to the phenotype, at least in older, cognitively-able individuals, who predominate in eye tracking studies conducted to date. Alternatively, it is possible that more complex stimuli make processing easier and more consistent by directing attention via movement or sound to key socially-relevant elements or that other cognitive deficits not explicitly evaluated by this literature (e.g. sensorimotor control) negatively impact the early development of appropriate social perception. Future studies are needed that evaluate and compare these possibilities.

Overall effects were larger for nonsocial than social stimuli (with the exception of eyes and faces within human interaction stimuli), and this effect remained constant across age. The finding of larger overall effects for nonsocial stimuli is consistent with our recent study²² and suggests several nonexclusive possibilities: a) a basic discrimination problem in selecting the most salient stimuli, b) difficulty sustaining attention to the most salient stimuli and filtering extraneous information during visual perceptual experience, and/or c) a top-down attentional preference for stimuli that are not socially or emotionally relevant in most contexts but have acquired reinforcing value via prior experience (e.g. restricted interests and abnormal sensory interests). While the present meta-analysis cannot definitively distinguish between these possibilities, consistency of effects across age decreases the likelihood that reinforcement history is playing a major role. Furthermore, the fact that we have found abnormal attention to nonsocial stimuli within the first few seconds of presentation²² decreases the likelihood of primary sustained attention deficits. The most plausible explanation appears to be a basic discrimination problem that emerges early in life and remains relatively constant. We would note that this interpretation would also imply that higher-order social cognitive difficulties (problems with complex social inference and perspective taking) may be rooted in more basic social/nonsocial discrimination processes. For example, it may be that detecting subtle intentions in other people is negatively impacted by a tendency to be flooded with relevant and extraneous information resulting in difficult sorting the key pieces and making sense of them in real time.³¹ Another plausible explanation is that autism results from variable dysfunction in the social circuitry important for theory of mind, mentalizing, and emotional salience,^{32,33} with the level of circuit dysfunction dictating whether cognitive deficits are restricted to higher-order processes or more basic social attention and discrimination. Future research is needed to evaluate these possibilities and to examine whether a dysfunctional preference for nonsocial stimuli is present in first-degree family members³⁴ or resolves in children with optimal outcomes.³⁵ This work will be crucial for determining whether increased gaze to non-social information tracks closely with overt symptomatology or represent an underlying cognitive diathesis.

Methodological factors strongly moderated effect sizes, reinforcing the need for future studies to use and report the most rigorous eye tracking methods, including methodological factors such as visual angle, ROIs size, and how missing data was handled across study groups. Points of calibration, the number of stimuli presented, and the type of control comparison appear to be key factors that can be addressed in study design. Importantly, though, even in high rigor studies, effects for human interaction stimuli remained highly significant and produced medium-to-large effect sizes. Future investigations should ensure that participants and gaze tracking parameters are well characterized and clearly reported. Studies should be particularly attentive to the number of calibration points used, as this may increase calibration accuracy, and the number of stimuli presented. These factors are easily controlled and may increase statistical power. The present review also identified that reporting of methodology was historically quite weak, although it is important to note that reporting may not always reflect actual methodology. Improvements in this area are still needed and will likely force the field toward higher standards in research design, increasing the likelihood of replication.

Effect sizes were remarkably constant across the wide range of ages sampled. The lack of worsening or improvement across ages implies that social attention abnormalities develop very early in life and remain throughout adulthood. This pattern is consistent with findings of persistent abnormal functional connectivity in key social perception networks in adolescents.³⁶ Similarly, the proportion of males and IQ levels for each study—as well as IQ differences between autism and control groups—did not moderate effect magnitude. Effect consistency across sex distributions is less surprising because most samples had high proportions of males and meta-analytic tests do not directly examine within-study sex differences for specific stimuli. However, stable effect magnitude across IQ levels and IQ differences further supports the core nature of gaze abnormalities to the autism phenotype. At a practical level, this also implies that matching on these factors, while generally useful, may be less crucial in this context. The specific choice of gaze measure also seems to be less important—all metrics had roughly similar average effect sizes.

A trend-level tendency was observed toward smaller effects when patients with autism were compared to DD rather than healthy controls. This trend is important in two ways. First, it suggests residual gaze abnormalities in other developmental conditions. Understanding these residual abnormalities may be important for characterizing the social problems that often manifest in these conditions. Second, the residual gaze abnormalities observed in DD controls were substantially smaller than those seen in autism. As a result, gaze measurements to social and non-social stimuli may differentiate autism from other neuropsychiatric conditions. Specifically, medium-sized effects for nonsocial, eye, and whole face ROIs within human interaction stimuli, if aggregated, could produce sufficient validity to inform autism identification.²²

It is also important to note that a clear publication bias was observed, with early studies of lower quality being disproportionately represented. However, the impact of these studies was reduced when removing outliers and accounting for points of calibration. Analyses of high rigor studies supported significant and meaningful effect sizes. Power to detect small modifier effects was low, but power was sufficient for detecting the major hypotheses regarding stimulus and ROI influences. As additional studies are added to this literature, this review can be updated and added studies leveraged to increase sensitivity to smaller moderator effects. The present review was intentionally limited to gaze duration and time-to-first-fixation measures. Including other gaze measures (saccades, blink, etc.) in future reviews may provide additional specificity to the pattern and further clarify the nature of the core cognitive features of autism. While dynamic stimuli did not increase effect magnitude, including pattern analysis in the context of these stimuli will be important for understanding temporal and spatial aspects of attention in autism. Similarly, incorporating pattern analysis with other gaze measures in the context of dynamic stimuli could also improve autism-control group separation, producing more powerful autism-specific measures for risk assessment and outcome tracking.

In sum, though exaggerated in earlier and less rigorous studies, gaze abnormalities in autism are not a myth. In this comprehensive meta-analytic investigation of attention to social and non-social stimuli, gaze abnormalities consistent with social cognitive deficits were identified as a reliable feature of ASD. These findings suggest ways to improve the rigor and

power of future investigations, highlight the need to develop and test specific social interventions, and provide justification for the development of objective, quantitative, gaze-based risk and outcome measures to advance clinical evaluation autism.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was made possible by a generous donation from the Stephan and Allison Cole Family Research Fund. The work was also supported by funding from the Case Western Reserve University International Center for Autism Research and Education (ICARE) and funding for the Developmental Synaptopathies Consortium (U54NS092090). The Developmental Synaptopathies Consortium is part of NCATS Rare Disease Clinical Research Network (RDCRN), an initiative of the Office of Rare Disease Research (ORDR). This consortium is funded through collaboration between NCATS, and the National Institute of Neurological Disorders and Stroke (NINDS) of the National Institutes of Health. The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Dr. Eng is the Sondra J. and Stephen R. Hardis Endowed Chair of Cancer Genomic Medicine at the Cleveland Clinic, and an American Cancer Society Clinical Research Professor.

TWF designed the present study. TWF obtained funding to support article review, coding, and analyses. TWF, EWK, and EEZ supervised the article review process. TWF, MSS, AYH, and EAY supervised interpretation of the study. TWF conducted data management and data analyses. All authors contributed to writing and revision.

This study was submitted for presentation to the International Meeting for Autism Research in San Francisco, CA May 10–13, 2017.

Drs. Frazier and Youngstrom served as the statistical experts for this research.

Dr. Frazier has received federal funding or research support from, acted as a consultant to, received travel support from, and/or received a speaker's honorarium from the Cole Family Research Fund, Simons Foundation, Ingalls Foundation, Forest Laboratories, Ecoeos, Curemark, IntegraGen, Kugona LLC, Shire Development, Bristol-Myers Squibb, National Institutes of Health, and the Brain and Behavior Research Foundation. Dr. Hardan has received research funding from Edison Pharmaceuticals and has served as a consultant to Hofman Technologies and Roche. Dr. Youngstrom has received grant or research support from the National Institute of Mental Health and has received internal funds from UNC-Chapel Hill. He has served on the advisory board/DSMB of the International Bipolar Foundation. He has served as a consultant about assessment to Joe Startup Technologies, Otsuka, Pearson Publishing, Western Psychological Services, Janssen, and Lundbeck. Mr. Klingemier has received research support from Kugona LLC.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Fifth. Arlington, VA: American Psychiatric Association; 2013.
2. Kanner L. Autistic disturbances of affective contact. *Nervous Child*. 1943; 2:217–250.
3. Rutter, M., Le Couteur, A., Lord, C. Autism Diagnostic Interview-Revised Manual. Los Angeles: Western Psychological Services; 2003.
4. Lord, C., Rutter, M., DiLavore, PC., Risi, S., Gotham, K., Bishop, SL. Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part 1): Modules 1–4. Torrance, CA: Western Psychological Services; 2012.
5. Jones W, Klin A. Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism. *Nature*. 2013; 504(7480):427–431. [PubMed: 24196715]
6. Sasson NJ, Elison JT. Eye tracking young children with autism. *J Vis Exp*. 2012; (61):3675. pii. [PubMed: 22491039]
7. Papagiannopoulou EA, Chitty KM, Hermens DF, Hickie IB, Lagopoulos J. A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders. *Soc Neurosci*. 2014; 9:610–632. [PubMed: 24988218]

8. Chevallier C, Parish-Morris J, McVey A, et al. Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Res.* 2015; 8:620–8. [PubMed: 26069030]
9. Chawarska K, Macari S, Shic F. Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. *Biol Psychiatry.* 2013; 74(3):195–203. [PubMed: 23313640]
10. Rice K, Moriuchi JM, Jones W, Klin A. Parsing heterogeneity in autism spectrum disorders: visual scanning of dynamic social scenes in school-aged children. *J Am Acad Child Adolesc Psychiatry.* 2012; 51(3):238–248. [PubMed: 22365460]
11. Chita-Tegmark M. Attention allocation in ASD: A review and meta-analysis of eye-tracking studies. *Review Journal of Autism and Developmental Disorders.* 2016; 3:209.
12. Chita-Tegmark M. Social attention in ASD: A review and meta-analysis of eye-tracking studies. *Res Dev Disabil.* 2016; 48:79–93. [PubMed: 26547134]
13. Nelson CA 3rd. Facing autism: eye tracking paves the way to underlying mechanisms. *J Am Acad Child Adolesc Psychiatry.* 2012; 51(3):233–234. [PubMed: 22365458]
14. Ornitz EM. The modulation of sensory input and motor output in autistic children. *J Autism Child Schizophr.* 1974; 4(3):197–215. [PubMed: 4374459]
15. Marco EJ, Hinkley LB, Hill SS, Nagarajan SS. Sensory processing in autism: a review of neurophysiologic findings. *Pediatr Res.* 2011; 69(5 Pt 2):48R–54R.
16. Chevallier C, Kohls G, Troiani V, Brodtkin ES, Schultz RT. The social motivation theory of autism. *Trends Cogn Sci.* 2012; 16(4):231–239. [PubMed: 22425667]
17. Baron-Cohen S, Leslie AM, Frith U. Does the autistic child have a “theory of mind”? *Cognition.* 1985; 21(1):37–46. [PubMed: 2934210]
18. Gallese V. Intentional attunement: a neurophysiological perspective on social cognition and its disruption in autism. *Brain Res.* 2006; 1079(1):15–24. [PubMed: 16680812]
19. Ramachandran VS, Oberman LM. Broken mirrors: a theory of autism. *Sci Am.* 2006; 295(5):62–69.
20. Ozonoff S, Jensen J. Brief report: specific executive function profiles in three neurodevelopmental disorders. *J Autism Dev Disord.* 1999; 29(2):171–177. [PubMed: 10382139]
21. Frith U, Happe F. Autism: beyond “theory of mind”. *Cognition.* 1994; 50(1–3):115–132. [PubMed: 8039356]
22. Frazier TW, Klingemier EW, Beukemann M, et al. Development of an objective autism risk index using remote eye tracking. *J Am Acad Child Adolesc Psychiatry.* 2016; 55(4):301–309. [PubMed: 27015721]
23. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med.* 2009; 151:264–269. W264. [PubMed: 19622511]
24. Lipsey, MW., Wilson, DB. *Practical Meta-analysis.* Thousand Oaks, CA: Sage Publications; 2001.
25. Kowatch RA, Youngstrom EA, Danielyan A, Findling RL. Review and meta-analysis of the phenomenology and clinical characteristics of mania in children and adolescents. *Bipolar Disorders.* 2005; 7:483–496. [PubMed: 16403174]
26. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw.* 2010; 36:1–48.
27. *Metafor: meta-analysis package for R* [computer program]. Version 1.9-9 2016.
28. Publications APA, Communications Board Working Group on Journal Article Reporting S. Reporting standards for research in psychology: why do we need them? What might they be? *Am Psychol.* 2008; 63:839–851. [PubMed: 19086746]
29. Hedges LV, Pigott TD. The power of statistical tests in meta-analysis. *Psychol Methods.* 2001; 6:203–217. [PubMed: 11570228]
30. Pellicano E, Maybery M, Durkin K, Maley A. Multiple cognitive capabilities/deficits in children with an autism spectrum disorder. “Weak” central coherence and its relationships to theory of mind and executive control. *Dev Psychopathol.* 2006; 18:77–98. [PubMed: 16478553]

31. 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. [PubMed: 11079426]
32. Di Martino A, Shehzad Z, Kelly C, et al. Relationship between cingulo-insular functional connectivity and autistic traits in neurotypical adults. *Am J Psychiatry*. 2009; 166:891–899. [PubMed: 19605539]
33. Baron-Cohen S, Ring HA, Bullmore ET, Wheelwright S, Ashwin C, Williams SC. The amygdala theory of autism. *Neurosci Biobehav Rev*. 2000; 24:355–364. [PubMed: 10781695]
34. Dalton KM, Nacewicz BM, Alexander AL, Davidson RJ. Gaze-fixation, brain activation, and amygdala volume in unaffected siblings of individuals with autism. *Biol Psychiatry*. 2007; 61:512–20. [PubMed: 17069771]
35. Fein D, Barton M, Eigsti IM, et al. Optimal outcome in individuals with a history of autism. *J Child Psychol Psychiatry*. 2013; 54:195–205. [PubMed: 23320807]
36. Fishman I, Keown CL, Lincoln AJ, Pineda JA, Muller RA. Atypical cross talk between mentalizing and mirror neuron networks in autism spectrum disorder. *JAMA Psychiatry*. 2014; 71:751–760. [PubMed: 24740586]

Clinical Guidance

- Gaze deficits in autism are apparent in basic looking paradigms, supporting inclusion of social attention as a key symptom within assessment instruments and diagnostic criteria.
- The magnitude of looking differences in autism is large enough that, if combined across regions and stimuli, this information may be useful for creating objective tools to enhance case identification.
- Age, sex, and cognitive level did not moderate findings, suggesting that gaze deficits may be useful for diagnosis and treatment tracking across the full spectrum of at-risk individuals.

Figure 1.
Flow diagram of search strategy, inclusion/exclusion, and final sample.

Figure 2.

Stronger methodology produces lower effect estimates (A–C), but effects remain significant with medium effect sizes under high rigor designs (D). Note: DD = developmental disability.

Figure 3.

Nonsocial regions-of-interest (ROI) show larger effect sizes than social ROIs, and the effect is strongest for human interactive stimuli (A); eyes and whole face ROIs show the largest effect sizes (B). Note. P-value represents significantly large effect sizes for social and nonsocial ROIs contained within human/interactive stimuli.

Table 1
Multivariate Mixed Effects Meta-Regression Estimates of the Effects of Sample Characteristics and Study Methodology.

Method Model (k=1155)					
	b	SE	z	p	95% CI
Intercept	0.40	.07	5.56	<.001	.26 to .54
Age (autism group)	<0.01	.00	-0.71	.475	-.01 to .01
DD controls (vs. healthy)	<i>-0.11</i>	<i>.06</i>	<i>-1.92</i>	<i>.055</i>	<i>-.23 to .00</i>
Publication year (centered at 2011)	-0.02	.01	-2.32	.020	-.04 to -.01
Design and reporting quality index (centered at 8.63)	-0.07	.02	-4.64	<.001	-.10 to -.04
Gaze measure type (vs. fixation count)					
Fixation count percent	0.11	.09	1.26	.208	-.06 to .29
Fixation duration	0.02	.06	0.30	.761	-.10 to .13
Fixation duration percent	0.02	.06	0.35	.728	-.10 to .15
Time to first fixation	0.12	.08	1.46	.145	-.04 to .28
Individual Follow-Up Models (k=796 to 1140)					
	b	SE	z	p	95% CI
Sex (autism group %)	-0.01	.01	-1.13	.260	-0.01 to .01
IQ (autism group mean)	-0.01	.01	-0.54	.589	-0.07 to .01
IQ difference (control - autism)	-0.01	.01	-0.74	.459	-0.01 to .01
Sampling rate	0.03	.03	0.99	.322	-0.03 to .09
Points of calibration	0.16	.05	3.19	.001	0.06 to .25
Number of stimuli	<i>-0.03</i>	<i>.02</i>	<i>-1.74</i>	<i>.082</i>	<i>-0.07 to .01</i>

Note. Bold designates factors significant at $p < .05$. Italics indicate factors that are trending at $p < .100$ but $p > .05$. Sex $k=1049$, IQ $k=942$, IQ Difference $k=852$, Sampling Rate $k=1091$, Points of Calibration $k=796$, Number of Stimuli $k=1140$. DD=Developmental Disability.

Table 2

Multivariate Mixed Effects Meta-Regression Estimates of the Effects of Stimulus and Region-of-Interest (ROI) Moderators.

Simultaneous Moderator Model (k=1155)					
	b	SE	z	p	95% CI
Method Factors					
Intercept	0.27	0.11	2.45	.015	.05 to .49
Publication year (centered at 2011)	−0.02 [*]	0.01 [*]	−2.26 [*]	.024 [*]	−.04 to −.01 [*]
Design and reporting quality index (centered at 8.63)	−0.06 [*]	0.02 [*]	−4.06 [*]	<.001 [*]	−.09 to −.03 [*]
DD controls (vs. healthy)	−0.11 [*]	0.06 [*]	−2.00 [*]	.045 [*]	−.22 to −.01 [*]
Stimulus Factors					
Audio-visual (vs. visual only)	0.02	0.08	0.24	.810	−.13 to .17
Human	0.03	0.08	0.39	.699	−.13 to .19
Dynamic (vs. static)	0.23	0.17	1.41	.159	−.09 to .56
Face only (vs. other)	0.01	0.06	0.17	.866	−.13 to .14
Interactive	−0.28	0.18	−1.57	.117	−.63 to .07
Directive (vs. no directive)	0.02	0.05	0.41	.684	−.08 to .13
Dynamic X interactive	−0.13	0.28	−0.47	.637	−.68 to .42
Dynamic X human	−0.19	0.17	−1.12	.264	−.53 to .15
Human X interactive	0.43 [*]	0.18 [*]	2.31 [*]	.021 [*]	.07 to .79 [*]
Dynamic X interactive X human	0.07	0.30	0.22	.826	−.52 to .65
ROI					
Social (vs. nonsocial)	−0.11 [*]	0.05 [*]	−2.39 [*]	.017 [*]	−.20 to −.02 [*]
Body regions (vs. other face)					
Eyes	0.26 [*]	0.09 [*]	2.82 [*]	.005 [*]	.08 to .44 [*]
Whole face	0.29 [*]	0.09 [*]	3.12 [*]	.002 [*]	.11 to .47 [*]
Upper face	−0.14	0.17	−0.82	.415	−.48 to .20
Lower face	0.15	0.17	0.90	.367	−.18 to .49
Face ratio (upper/lower)	−0.13	0.15	−0.87	.387	−.44 to .17
Mouth	−0.01	0.09	−0.15	.879	−.20 to .17
Nose	−0.13	0.12	−1.08	.278	−.36 to .10

Simultaneous Moderator Model (k=1155)					
	b	SE	z	p	95% CI
Other body	−0.02	0.11	−0.22	.829	−.24 to .19
Non-body	0.10	0.09	1.17	.240	−.07 to .27

Note. DD=Developmental Disability.

* p<.05.