Neuroscience and Biobehavioral Reviews 179 (2025) 106391

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

The use of eye-tracking to find objective outcome measures of early intervention strategies for children with autism: A systematic review

Chloé Peter [a,*], Maria-Paraskevi Antoniou [a], Evelyne Antonietti [a], Joana Almeida Osório [a], Joëlle Rosselet Amoussou [b], Nadia Chabane [a,1], Borja Rodríguez-Herreros [a,1]

[a] Autism Spectrum Disorders Service, Department of Psychiatry, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland
[b] Medical Library-Cery, Site de Cery, Lausanne University Hospital and University of Lausanne, Prilly, Switzerland

ARTICLE INFO

Keywords:
Autism
Early intervention
Eye-tracking
RCTs
Visual attention

ABSTRACT

Objective: Early intervention strategies are recommended for children with autism to improve cognitive and social communication skills. However, there is a persistent challenge to identify objective outcome measures of intervention efficacy. Eye-tracking (ET) is a safe and well-tolerated technology able to detect differences in visual attention through gaze behavior. This systematic review aims to identify, appraise and summarize the existing literature using ET to track response to early intervention in autism clinical trials.
Method: A comprehensive literature search was conducted in December 2024 in several bibliographic databases, including Medline ALL Ovid, Embase.com, APA PsycInfo Ovid, the Cochrane Library Wiley and PubMed Central. Additionally, we searched trial registries and we performed citation tracking strategies. We retained randomized controlled trials (RCTs) incorporating an ET outcome measure to evaluate the effect of direct, parent-mediated and pharmacological early interventions. Quality was assessed using Risk of Bias 2 and GRADE.
Results: Out of 1'726 reports screened, eleven articles from nine unique RCTs met inclusion criteria, with overall low-to-moderate risk of bias. Six out of the nine intervention approaches showed improvements in several indices of visual attention measured with different ET paradigms. We also identified 24 RCTs registered with an ET outcome measure yet without published results.
Conclusion: ET proved to be a clinically relevant measure sensitive to change in several interventional contexts. However, not all ET paradigms measured change over time reliably, compromising interpretability. Further research is needed to ensure that ET potentially becomes an accessible and accurate tool for widespread adoption in clinical practice.

1. Introduction

Autism is a common and heterogeneous neurodevelopmental condition characterized by persistent and significant atypicalities in social communication and interaction, as well as the presence of restricted, repetitive and sensory behaviors (American Psychiatric Association, 2022). Recent epidemiological studies estimate that autism currently concerns approximately 78 million individuals worldwide (Lord et al., 2022; Zeidan et al., 2022). A diagnosis of autism is based exclusively on combining direct behavioral and clinical observations of the individual with indirect measures, such as a comprehensive developmental and medical history (Risi et al., 2006). A stable diagnosis is typically established around 24 months of age, when trained clinicians can reliably differentiate the above set of core diagnostic features from other neurodevelopmental trajectories (Ozonoff et al., 2015). There is a broad consensus on current international guidelines to advocate for early screening and diagnosis of autism, with the aim of enabling timely intervention and parental support (NICE, 2021).

Early intervention approaches encompass parent-led and/or therapist-administered programs during childhood. Medication may also be indicated to manage certain symptoms associated with autism, such as severely disruptive behaviors, irritability and co-occurring comorbidities like anxiety, rather than autism core symptomatology (Accordino et al., 2016; Hirota and King, 2023). In young children, intervention strategies are introduced as early as possible in order to capitalize on heightened brain plasticity, although more empirical

* Correspondence to: Service des Troubles du Spectre de l'Autisme & apparentés, Centre Cantonal de l'Autisme, Av. de Beaumont 23, Lausanne 1011, Switzerland
E-mail address: chloe.peter@chuv.ch (C. Peter).
[1] These authors contributed equally to this work

https://doi.org/10.1016/j.neubiorev.2025.106391
Received 20 March 2025; Received in revised form 18 September 2025; Accepted 30 September 2025
Available online 4 October 2025

evidence is warranted (Dawson, 2008). Several rigorous systematic reviews evaluating early intensive interventions for children with autism have shown moderate improvements in social communication competences, language skills and joint attention, potentially fostering further development and the reduction of symptom severity (Crank et al., 2021; Sandbank et al., 2020; Zhuang et al., 2024). Nevertheless, response to early intervention is variable and assessing its efficacy remains challenging due to the absence of appropriate tools (Kasari, 2002; Lord et al., 2005). A suitable outcome measure for evaluating response to intervention needs to be feasible to implement, blindly administered, objectively measured, clinically relevant and sensitive to change. Existing instruments are neither optimally designed to measure change nor sufficiently sensitive to individual variability, making comparison between studies difficult and hindering the development of standardized recommendations (Grzadzinski et al., 2020).

The identification of a biomarker –defined as an objective characteristic that can indicate abnormal biological processes and/or responses to therapeutic intervention (Atkinson Jr et al., 2001)–, would represent a breakthrough in autism research, both for refining diagnosis and for evaluating intervention outcomes (Califf, 2018; McPartland, 2016). To date, no reliable (i.e., appropriate sensitivity and specificity) biomarker has yet emerged as an indicator of change that can be integrated into clinical practice (Javitt and McPartland, 2024; Jensen et al., 2022). Examples of proposed response biomarkers are based on technologies that can be used to identify physiologically relevant abnormalities associated with autism, such as genetic, metabolomic, neuroimaging, neurophysiological and neurobehavioral (e.g., eye-tracking) measures (Parellada et al., 2023; Walsh et al., 2011). The latest are receiving particular attention due to their potential to provide standardized, cost-effective and noninvasive treatment outcome evaluation (McPartland et al., 2020; Shic, 2016). Specifically, eye-tracking (ET) is a safe, affordable and well-tolerated technology that provides an easily detectable signature of early differences in visual attention by the quantification of low-level behaviors such as eye movements (Chita-Tegmark, 2016). Below we will review key evidence for the potential of ET to provide insights into cognitive and attentional aspects of information processing, as well as an objective outcome measure to evaluate early intervention programs for young children with autism.

1.1. Eye-tracking as a window into cognitive processing

The rationale for using ET to capture intervention effects is grounded in well-established theoretical frameworks. The eye-mind hypothesis (EMH) (Just and Carpenter, 1980) has been foundational for interpreting gaze behavior as an online marker of attention and information uptake as they unfold in real time. This principle posits that the eyes remain fixated on a stimulus for as long as it is being processed, thereby providing a temporally precise index of comprehension, attention and behavioral measures of cognition (Liversedge and Findlay, 2000). When the eyes pause on a given location (fixation), visual information is encoded. EMH therefore assumes that fixation duration is a valid measure of the depth of cognitive processing. While influential and widely applied in ET studies, EMH is not absolute; alternative models highlight task-dependent modulation of eye–mind coupling (Rayner, 1998), such as parafoveal and peripheral processing in which covert attention (i.e., processing things not directly looked at) and task complexity could disrupt this direct relationship (Posner, 1980). These refinements highlight that the eye–mind relationship is probabilistic and task-dependent, but they do not undermine the core principle that ET provides a faithful window into cognitive processing, supporting its utility as a meaningful outcome measure.

ET also affords a rich set of parameters that offer converging, fine-grained insights into visual attention and information processing. A common approach is the study of eye fixations into pre-defined regions within a given stimulus, so-called areas of interest (AOIs), such as the eyes, the mouth, or a particular object in a scene (Holmqvist et al., 2011). Within each AOI, dwell time quantifies the total gaze duration, providing an estimate of the share of attention allocated to that particular area. Moreover, the number of discrete fixations (i.e., fixation count or rate) and its duration (i.e., the time that gaze remains still) within an AOI complements this information by indicating how thoroughly a region was explored (Henderson and Hollingworth, 1998). These periods of stable gaze are punctuated by saccades –rapid shifts between fixations–, which reveal how attention is dynamically redistributed across a scene (Salvucci and Goldberg, 2000). Taken together, these measures underscore the versatility of ET as a sensitive and interpretable tool for evaluating changes in the underlying mechanisms of information processing (Hessels et al., 2024; Rahal and Fiedler, 2019).

1.2. Eye-tracking in autism research

From birth, eye contact plays a crucial role in human nonverbal communication and interaction, and has been the subject of extensive research across multiple disciplines (Jongerius et al., 2020). Typically developing (TD) infants normally exhibit very early in their development a visual preference towards social cues in their environment, such as eyes, faces and voices (i.e., social preference) (Jones and Klin, 2013). The ability to visually explore an environment and give preferential attention to social stimuli shapes infant development and underlies many forms of interrelated learning, particularly imitation and joint attention (Aslin, 2012; Frank et al., 2009). These prerequisites exert a considerable influence on the development of social and language skills, and they seem to be affected in young children with autism (Dawson et al., 1998). Notably, difficulties when orienting to social cues and sharing attention may trigger a cascading developmental effect leading to difficulties in social communication (Campbell et al., 2014; Gliga et al., 2015; Jones et al., 2008). One way to measure visual attentional patterns is using ET, which provides an objective gaze assessment to investigate atypical neurodevelopmental trajectories in children with autism (Frank et al., 2009). Several studies have provided substantial evidence of disparities in how autistic children process visual information compared to TD children. For example, Shic and colleagues reported that autistic children exhibited significantly less sustained attention (i.e., mean look duration) to faces during socially engaging contexts (Shic et al., 2023). Likewise, a robust meta-analysis showed an overall reduction of attention to social stimuli, particularly in scenes involving multiple people (Chita-Tegmark, 2016). Converging evidence further indicates that autistic children show atypical patterns of face scanning (i.e., the distribution of gaze across facial features), specifically reduced fixation rates to the eye region (Falck-Ytter et al., 2013; Frazier et al., 2017; Klin et al., 2002; Papagiannopoulou et al., 2014), as well as a heightened preference for nonsocial (e.g., geometrical) cues (Pierce et al., 2016). These differences also extend to attentional flexibility, that is, the capacity to alternate between stimuli. While TD infants typically become faster at disengaging their gaze from one stimulus to orient toward another, infants later diagnosed with autism show prolonged latencies and reduced developmental improvements over time. This has been demonstrated using the gap-overlap paradigm, in which a peripheral target must be fixated while a central stimulus remains visible, revealing specific difficulties in attention disengagement (i.e., the ability to shift gaze away from a stimulus) (Elsabbagh et al., 2013).

Furthermore, the convergent validity of ET indices of social attention has been corroborated by their robust association with caregiver ratings on clinical measures assessing social communication (Murias et al., 2018). Such findings have supported the use of ET to provide a potential diagnostic biomarker for autism (Chita-Tegmark, 2016; Papagiannopoulou et al., 2014). With its ability to capture subtle gaze and attentional patterns in young children (Shic et al., 2022), the utility of ET has expanded to become an early detection tool, in particular because distinctive patterns of eye movements already emerge in toddlers at risk of autism long before behavioral symptoms become apparent (Hou et al., 2024a; Jones and Klin, 2013). Recently, machine-learning (ML) and

deep-learning models have garnered significant attention due to their high accuracy and efficiency when classifying individuals with and without autism based on their gaze patterns measured with ET (Alsharif et al., 2024; Asmetha Jeyarani and Senthilkumar, 2023; Kanhirakadavath and Chandran, 2022). Such advancements pave the way to make early detection more accessible, particularly in primary care settings where traditional diagnostic resources may be limited.

The extensive study of ET as an early diagnostic and detection instrument in autism finds no parallel in its use to provide a neurobehavioral biomarker of response to intervention. There has been a remarkable increase in the number and quality of randomized controlled trials (RCTs) conducted to assess the efficacy of early intervention strategies in autism (Green and Garg, 2018). Nonetheless, ET has been scarcely incorporated to elucidate the neurobehavioral mechanisms that may circumvent clinically observed developmental changes (Anagnostou et al., 2015; Provenzani et al., 2020). Some proof-of-concept evidence has already shown that, with the appropriate choice of stimuli, ET could make a meaningful contribution to the array of outcome measures available to test psychosocial interventions in autism (Fletcher-Watson and Hampton, 2018). The use of ET as an outcome measure could therefore enhance our understanding of neural-to-behavior mechanisms underlying autism core symptoms. For instance, Green et al. (2015) conducted an RCT to evaluate the efficacy of a parent-mediated intervention (PMI) for infants at high risk of autism, incorporating ET as an outcome measure to assess changes in attention disengagement. This study exemplifies the type of design central to our review: rigorous early intervention trials in young autistic or high-risk children that employ ET to capture intervention-related changes. However, there has been a paucity of systematic reviews appraising the use of neurobehavioral outcome measures in autism clinical trials (McGlade et al., 2023). The aim of this systematic review is to examine the potential of ET to meaningfully capture the effectiveness of early interventions in young children with autism.

2. Method

The present systematic review has been elaborated following the guidelines and methodology of the Cochrane Handbook for Systematic Reviews of Interventions (Cochrane, 2024). The Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement was used for the reporting of the findings (Page et al., 2021). The review protocol was registered in the Prospective Register of Systematic Reviews (PROSPERO, CRD42025642433).

2.1. Information sources and search strategy

A literature search was conducted in December 2024 in collaboration with a medical librarian (JRA). The following bibliographic databases were consulted: Medline ALL Ovid, Embase.com, APA PsycInfo Ovid, the Cochrane Database of Systematic Reviews Wiley and the Cochrane Central Register of Controlled Trials Wiley. Searches combined the overarching concepts of “Autism”, “Eye-Tracking” and “Children”. For each concept we searched for free-text keywords and controlled vocabulary terms (e.g., MeSH). An RCT filter was applied. To capture studies where “eye-tracking” and synonyms appeared only in the full text, we also searched PubMed Central. The searches were performed without language or date restrictions. Deduplication of references exported from the databases was performed by Deduklick (Borissov et al., 2022). A backward and forward citation search was conducted on the retained studies using ResearchRabbit (Sharma et al., 2022); results were then exported to Endnote 20 (Clarivate Analytics, USA), and only references containing the terms (randomised OR randomized OR randomly) in "all fields" of the record, or the term (Trial) in the "title" were screened. A search targeting conference proceedings from the last three years (2022–2025) was also carried out in Embase.com. Finally, trial registry records were searched in the following registers: Cochrane Central Register of Controlled Trials Wiley, ClinicalTrials.gov and ICTRP (WHO International Clinical Trials Registry Platform). All final strategies are documented in the Supplemental Material, detailing the search syntax, keywords, index terms and filters used.

2.2. Eligibility criteria

Our inclusion criteria were structured using the PICO framework. **Population**: we included RCTs published between 2005 and 2025 that assessed the efficacy of early interventional approaches on children with autism or with an increased familial risk, whose age range started below 4 years and did not exceed 8 years of age. **Intervention**: we considered all early psychosocial and pharmacological interventions. We defined *psychosocial interventions* as “*interpersonal or informational activities, techniques, or strategies that target biological, behavioral, cognitive, emotional, interpersonal, social, or environmental factors with the aim of reducing symptoms of these disorders and improving functioning or well-being*” (England et al., 2015). Examples include naturalistic developmental behavioral interventions (e.g., Early Start Denver Model, ESDM), PMIs, or cognitive-behavioral strategies adapted for young children with autism. *Pharmacological interventions* referred to any study evaluating the administration of medication with the purpose of improving core or associated symptoms of autism. Interventional approaches that were outside these definitions were excluded. **Comparison**: Eligible comparators included alternative interventions, treatment-as-usual (TAU), or control group. **Outcome**: at least one measure obtained with ET, either as a primary or secondary outcome, was required.

Exclusion criteria: Studies were excluded if they 1) did not include a psychosocial or pharmacological intervention as defined above (Cheuk et al., 2011); 2) targeted age groups outside the specified range (Alvares et al., 2019); 3) targeted population without or at no risk of autism (Fletcher-Watson and Hampton, 2018); 4) lacked ET-based outcome measures (Brian et al., 2017); 5) were not RCTs (e.g., single-case studies, observational studies, or quasi-experimental designs) (Dawson et al., 2017; Latrèche et al., 2021); and 6) were not in English (Delrobaee et al., 2020).

2.3. Study selection

The database searches revealed 1’726 potentially relevant citations (Fig. 1). The search results were uploaded to Rayyan (Ouzzani et al., 2016) for the initial screening. The studies were independently screened by two authors (CP and BRH) based on title and abstract, resulting in 99 studies for which the full text was obtained. Full texts were read to ascertain whether the pre-selected studies met full inclusion criteria. Reasons for exclusion were reached by consensus between reviewers and documented. Where consensus was not reached, a third reviewer (MPA) resolved the disagreement. The most frequent reasons for exclusion were: 1) Wrong population (e.g., studies including only adolescents or adults with autism (van Pelt et al., 2020), or participants without a confirmed diagnosis or familial risk for autism) (Peters et al., 2021); 2) Wrong outcome (e.g., studies without an outcome measure obtained with ET) (Soriano et al., 2020); 3) Wrong study design (e.g., non-randomized, observational studies) (Murias et al., 2018; Vivanti et al., 2022); 4) Wrong publication type (e.g., reviews, commentaries) (Dawson et al., 2012; Hustyi et al., 2023). Additional citation tracking strategies and trial registries identified 382 supplementary reports. Two of these studies met full selection criteria, and we found 24 registered RCTs with ET unpublished data. Finally, 11 RCTs were retained.

2.4. Data extraction

Information extracted from each study included 1) general aspects of the RCT such as publication year, location and sample size; 2) demographic variables of the participants (age, sex, autism diagnosis); 3) type of intervention approach and comparator; 4) type of ET

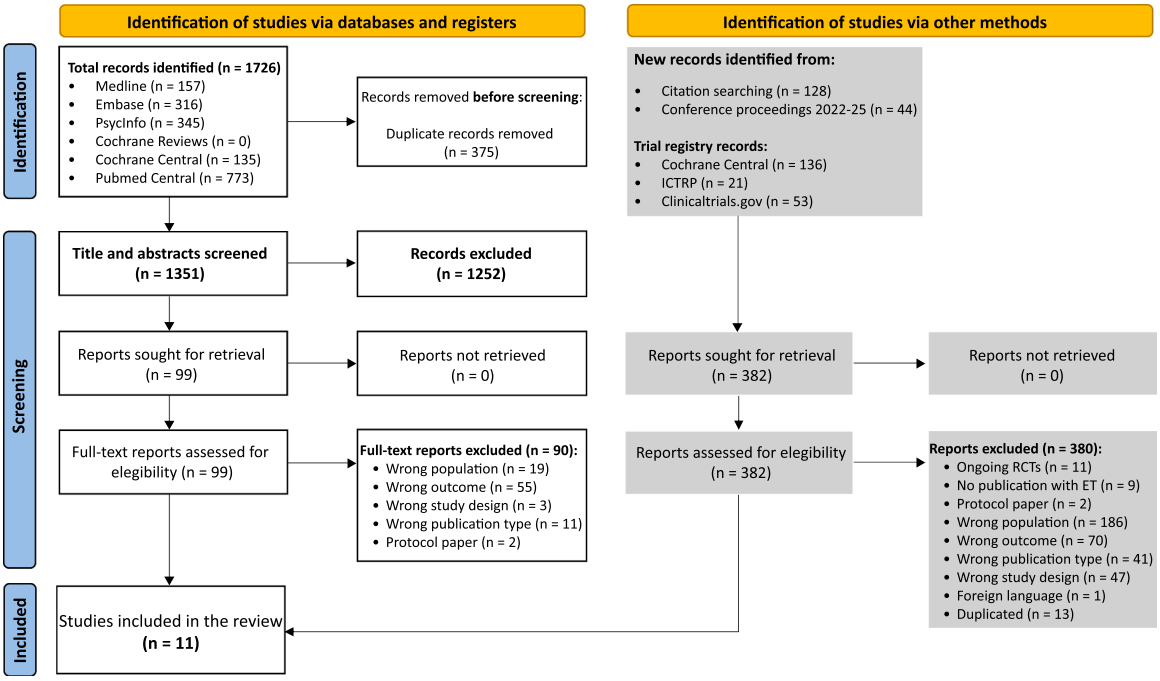

Fig. 1. PRISMA Flow Diagram (Page et al., 2021) illustrating study screening and selection process. Reasons for exclusion included wrong population (studies including children without autism, or only adolescents or adults with autism); wrong outcome (not obtained using ET); wrong study design (open-label trials, observational studies); wrong publication type (reviews, commentaries); protocol papers (without results); languages other than English and ongoing RCTs without published data. Examples of each exclusion are provided in the Methods section.

experimental paradigm; 5) ET outcome measure; 6) statistical design and 7) main findings. Due to the large heterogeneity of the retained RCTs, not only with respect to the intervention strategy but also the ET outcomes, only an exploratory meta-analysis was performed. A synthesis of the extracted data was undertaken narratively according to Cochrane guidelines (Higgins et al., 2024).

2.5. Risk of bias and GRADE assessment

The risk of bias of the eleven RCTs included in this review was evaluated by three reviewers (CP, MPA and BRH), using version 2 of the Cochrane Risk-of-bias tool for randomized trials (RoB-2) (Sterne et al., 2019). Each assessment was recorded within the Excel table provided by the RoB-2 tool with a brief rationale for each decision. The evaluation covered risk of bias arising from five key domains when considering the ET outcome measure of each study: i) randomization process, ii) deviation from the intended intervention, iii) missing outcome data, iv) measurement of outcomes, and v) selection of the reported results. The overall risk of bias was marked as high if any of these domains was rated as high, and it was marked as overall low if all domains were rated as low. In cases of uncertainty, the final determination was made through consensus among the three reviewers. Performance bias was overlooked on the basis that it was generally not possible to blind parents or trainers.

The certainty of evidence for each ET outcome measure was assessed using the GRADE (Grading of Recommendations, Assessment, Development, and Evaluations) approach (Balshem et al., 2011). We identified the main ET-derived outcomes of interest (e.g., attention disengagement, mouth-to-eyes ratio, social preference, and total fixation time to the eyes and to the mouth) and summarized the available evidence per outcome across the included studies. The certainty of the evidence could be labelled as either 'high,' 'moderate,' 'low,' or 'very low.' In line with GRADE methodology, RCTs were initially rated as high-certainty evidence. Certainty ratings were then downgraded based on concerns in five domains: risk of bias, inconsistency, indirectness, imprecision, and publication bias. Where applicable, ratings could also be upgraded (e.g., in cases of large effect sizes or dose–response gradients). For each outcome, a final certainty rating was assigned. Results were summarized in a GRADE "Summary of Findings" table, generated using GRADEpro GDT (GRADEpro, 2024).

2.6. Statistical analysis

Given the heterogeneity of ET metrics across studies, we conducted an exploratory meta-analysis restricted to outcomes reported by at least two RCTs, yielding the five abovementioned ET-derived measures. Standardized mean differences (SMD; Cohen's d) were calculated for each study to allow comparison across different ET metrics, and effect estimates were oriented such that positive values indicated beneficial intervention effects. Considering the methodological diversities among included trials, pooled estimates were computed using a random-effects model with restricted maximum likelihood (REML) estimation, which provides an unbiased estimate of the between-study variance (τ^2). Statistical heterogeneity was assessed in outcomes with more than three contributing studies using the Higgins I^2 statistic (percentage of total variation attributable to heterogeneity), and H^2 (ratio of between-study variability) (Higgins and Thompson, 2002). Cochran's Q statistic was also computed as an index of statistical heterogeneity. Where correlations between baseline and post-treatment values were unavailable, a conservative default of $r = 0.5$ was imputed (Chandler et al., 2019). For each pooled outcome, 95 % confidence intervals (CI) were generated, and forest plots were produced to visualize individual and summary effects. Overall, estimates were interpreted with caution given the limited precision and reliability of heterogeneity statistics under such conditions. All analyses were performed using R (v4.4.2) and the *metafor* package (Viechtbauer, 2010). Publication bias was explored through Egger's test and visual inspection of funnel plots (Egger et al., 1997). Results were synthesized alongside GRADE certainty ratings to provide a combined evaluation of effect sizes and strength of evidence.

3. Results

3.1. Study selection

The main characteristics of the eleven studies that met full inclusion criteria are summarized in Table 1. Despite the broadening of the search scope from 2005 to 2025, the eleven retained reports were all published within the last decade, thereby underscoring the contemporary relevance of the subject. Notably, only one RCT was not conducted in a Western country, with five of the studies being carried out in the United States. A higher proportion of young autistic boys compared to girls took part in all RCTs except for the three iBASIS-VIPP articles. The eleven published studies derived from nine unique RCTs, since the iBASIS-VIPP RCT featured two additional follow-up articles with the same participants, one at a later timepoint and another reporting an alternative ET outcome measure of social attention. The age of the participants ranged from 7 months to 8 years, with a mean age at inclusion in the study that varied from 22 to 68 months of age. The three iBASIS-VIPP studies were conducted with infants with an elevated familial likelihood for autism, whereas older children who had already received a diagnosis

Table 1
Overview of included studies.

Author, Year	Population	Age	Intervention & Comparator	ET paradigm	ET measure	Statistical design	Main Results
Green et al., (2015)	High risk of autism N = 54 (29 M, 25 F)	7–10 months	iBASIS-VIPP ***vs*** No Intervention	Gap-Overlap	Attention disengagement: saccadic reaction time	Regression analysis adjusted for baseline variables	Faster disengagement in the Gap-overlap task (d= 0·48, p = 0.077). **Favours iBASIS-VIPP**
Green et al., (2017)	High risk of autism N = 54 (29 M, 25 F)	7–10 months	iBASIS-VIPP ***vs*** No Intervention	Gap-Overlap	Attention disengagement: saccadic reaction time	Estimates from repeated measures within correlated regressions	No difference between groups at 27 months follow-up. Missing data at 39 months
Bedford et al., (2024)	High risk of autism N = 54 (29 M, 25 F)	7–10 months	iBASIS-VIPP ***vs*** No Intervention	Gaze following task	Social attention: proportion of dwell time to the referent	Regressions standardized beta estimates and SEs	Reduced dwell time to referred object (β = 0.32, p = 0.03) at 15-month treatment endpoint. **Favours No intervention**
Bradshaw et al., (2019)	ASD N = 28 (24 M, 4 F) TD N = 23 (14 M, 9 F)	18–48 months	PRISM ***vs*** TAU	Face scanning; social preference	Mouth-to-Eyes ratio; social vs geometric	ANCOVA	No significant differences in ET changes between the PRISM and TAU groups. **Favours TAU**
Wang et al., (2020)	ASD N = 35 (31 M, 4 F) TD N = 41 (19 M, 22 F)	34 months	GCET ***vs*** No Cue condition	Social looking	Attention to faces: % of face fixations	Linear mixed model	Increased fixation duration on faces in Cue Condition (p = 0.015). **Favours intervention**
Gepner et al., (2022)	ASD N = 23 (20 M, 3 F)	3–8 years	Slowed Audiovisual input ***vs*** Standard speech therapy	Face scanning	Total fixation time on AOIs (mouth, eyes, face, outside the face)	Linear mixed model	Increase in total fixation time on the mouth (p = 0.008) and eyes (p < 0.001). **Favours intervention**
Potter et al., (2019)	ASD N = 58 (46 M, 12 F)	2–6 years	Sertraline ***vs*** Placebo	Passive viewing eye-tracking (PVET)	Proportions of fixation duration to the target and distractor word	ANCOVA	No main effect of treatment group (p = 0.94), no main effect of time point (p = 0.79), and no interaction (p = 0.73). **Favours placebo**
Dawson et al., (2020)	ASD N = 180 (143 M, 37 F)	2–7 years	Cord Blood Infusion ***vs*** Placebo	Social preference and sustained attention	Time spent looking at actress vs toys and average look duration	ANCOVA	Greater odds of gazing at the toys 6 months after treatment (OR=1.43): no effect of treatment on social preference. **Favours placebo**
Le et al., (2022)	ASD N = 41 (38 M, 3 F)	3–8 years	Intranasal Oxytocin + social interaction ***vs*** Placebo	Social preference (GeoPref test)	Time spent viewing AOIs of emotional faces	ANCOVA	More time spent viewing social stimuli (p = 0.03); increased time viewing eyes of angry, happy and neutral faces (p < 0.03); decreased for fearful faces (p = 0.02). **Favours intervention**
Haworth et al. (2018)	ASD N = 18 (12 M,6 F)	24–42 months	Classroom-Based Intervention + caregiver training ***vs*** Caregiver training only	Gaze to the actor's face and anticipatory gaze to the goal location	Action anticipation, goal extraction, and AOI look duration	Paired T-test ANOVA	No change in action anticipation after intervention (p = 0.517). Significant decrease in goal extraction in the intervention group (p = 0.01). Increase in the ratio of AOI look to the face (p = 0.016). **Favours intervention**
Bast et al. (2025)	ASD N = 60 (50 M,10 F) TD N = 52 (36 M,16 F)	24–66 months	A-FFIP ***vs*** EIAU	Reactive joint attention (RJA)	RJA Likelihood (%) RJA duration	Linear mixed model	Increase in RJA likelihood at end-of-intervention (OR=1.52) and follow-up (OR=2.38) **Favours intervention**

M= male, F= female, ASD= autism spectrum disorder, GCET=Gaze-contingent eye-tracking training, iBASIS-VIPP= adapted Video Interaction to Promote Positive Parenting, PRISM = Pivotal Response Treatment for Social Motivation, TAU= treatment as usual, AOIs= areas of interest, A-FFIP = Frankfurt Early Intervention Program for ASD, EIAU: early intervention as usual

participated in the remaining studies. Three of the studies also included a group of TD children (Bast et al., 2025; Bradshaw et al., 2019; Wang et al., 2020). Three different types of early intervention strategies were trialed: interventions directly delivered to the child such as behavioral therapist-mediated programs and other nonpharmacological approaches (5/11), PMIs (3/11) and pharmacological interventions (3/11). Several ET paradigms were included as primary or secondary outcome measures. Eight studies reported on ET outcomes assessing social domains, such as attention to faces, face scanning, social preference and sustained attention, amongst others. The remaining three studies reported on nonsocial ET outcomes such as attention disengagement or passive viewing to measure receptive language.

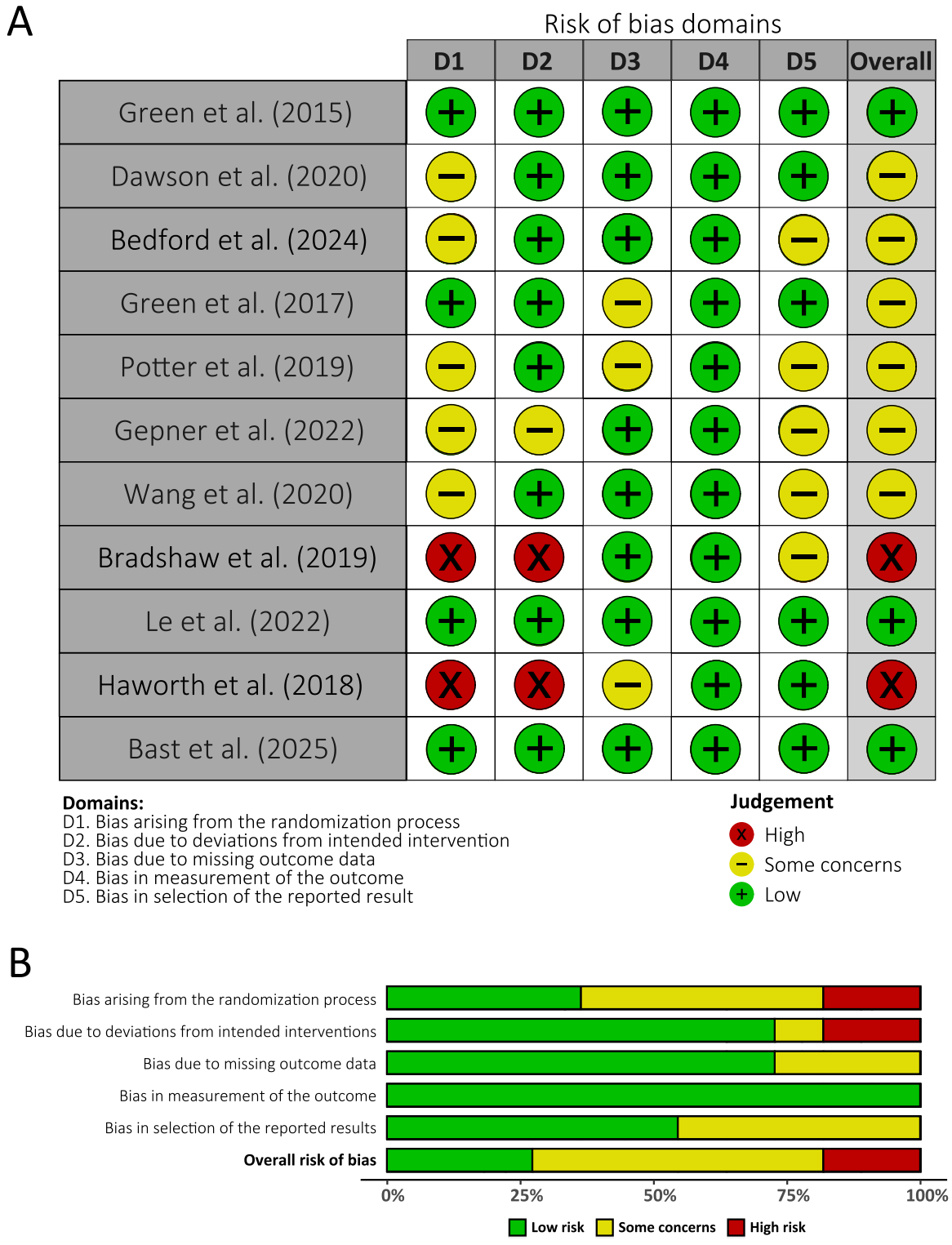

Fig. 2. Risk of bias summary of the included RCTs. The figure presents the evaluation of the risk of bias using the Cochrane Risk of Bias 2 (RoB-2) tool. Each domain is categorized as low risk, some concerns, or high risk of bias. The overall risk of bias is determined based on these domains.

3.2. Risk of bias assessment

Critical appraisal of the eleven studies is presented in Fig. 2. Two studies were judged as having an overall high risk of bias (Bradshaw et al., 2019; Haworth et al., 2018), specifically due to serious concerns with allocation concealment, differences at baseline that could impact the outcome measure, and no explicit mention of an intention-to-treat methodology. Six studies showed a moderate risk of bias, with five of them presenting some concerns relative to the randomization procedure, such as unclear allocation concealment (Gepner et al., 2022; Potter et al., 2019; Wang et al., 2020), baseline differences between the intervention groups (Dawson et al., 2020), or the lack of a baseline assessment of the ET outcome measure (Bedford et al., 2024). Only one of these six RCTs was judged to have potential deviations from the intended intervention (Gepner et al., 2022). Significant missing data was reported in the 27-month follow-up of the iBASIS-VIPP study, with up to 26 % (14/54) of participants lacking ET measures, and no data presented for the 39-month timepoint (Green et al., 2017). Similarly, the RCT evaluating the efficacy of sertraline could only report ET data from 14/58 participants (7 per group) (Potter et al., 2019). However, none of the eleven studies presented concerns regarding the measurement of the ET outcome measure. Two of the six studies with moderate risk of bias presented low concerns concerning selective reporting, as there was a prespecified analysis plan (Dawson et al., 2020; Green et al., 2017). The four remaining reports, together with one of the studies judged as having an overall high risk of bias, did not provide a pre-specified analysis for the ET outcome. Lastly, three studies (Bast et al., 2025; Green et al., 2015; Le et al., 2022) were categorized with low risk of bias across all achievable domains of the RoB-2 tool. All eleven reports stated that the funding institutions did not influence study methodology, data analysis or results reporting.

3.3. Results summary

In the following subsections, we provide an integrative summary of the effects reported on the ET outcome measures of the eleven retained studies. The main characteristics and findings for each report are highlighted according to the type of intervention delivered (PMIs, direct, and pharmacological).

Parent-mediated interventions(PMIs). Three interrelated studies were conducted to examine the impact of iBASIS-VIPP, a PMI designed to improve social communication in infants at elevated familial risk for autism. Across these studies, ET outcome measures revealed mixed findings. The primary study (Green et al., 2015) incorporated an ET outcome to explore changes in attention disengagement, by evaluating how quickly infants shifted attention to a peripheral target using the gap-overlap paradigm (Elsabbagh et al., 2013). Infants at risk of autism receiving the intervention showed marginally faster disengagement times compared to those undergoing no PMI. The authors suggest that promoting caregiver-infant social interaction enhanced nonsocial attentional flexibility. The anticipated moderate-sized positive effect, with an average decrease of 50 ms in attention disengagement, was very close to statistical significance. Nevertheless, a later follow-up assessment (Green et al., 2017) at 27 and 39 months to ascertain long-term consolidation effects of the intervention reported a lower difference compared to baseline, and no significant difference between groups. The absence of ET data at 39 months restricts the possibility of drawing any definitive conclusion about potential consolidated effects on nonsocial attentional flexibility over time. A third study shifted focus to dwell time on objects referenced by an adult (Bedford et al., 2024). Counterintuitively, infants in the iBASIS-VIPP group displayed significantly reduced dwell times compared with the control group. The authors claim that this reduction may reflect efficient attention allocation, potentially allowing for faster engagement and disengagement with social cues. The interpretation of these results is called into question given conventional associations between prolonged fixation and superior joint attention. Taken together, these findings suggest that PMIs may influence attentional mechanisms of infants at elevated risk for autism, particularly disengagement, indicating a potential to shape early developmental trajectories. However, methodological differences (e.g., paradigms used, timing of assessments, analytic choices) likely makes evidence inconsistent and difficult to reconcile across studies. More broadly, the studies illustrate both the promise and the current limitations of ET in capturing subtle intervention-related changes in early attentional patterns.

Direct interventions. Five RCTs used ET as an outcome to measure the impact of early interventions directly delivered on young autistic children. Results were heterogeneous across paradigms. A pilot RCT evaluating the Pivotal Response Treatment for Social Motivation (PRISM) compared to TAU found no group differences in social preference or mouth-to-eye ratio. However, the study replicated lower preference for social stimuli in autistic children compared to TD participants (Pierce et al., 2011), reinforcing social preference as a potential diagnostic biomarker but suggesting it may lack sensitivity to detect intervention-related changes. By contrast, another RCT examining the effects of slowed audiovisual input compared to standard speech therapy reported a higher rate of face fixations in both groups, with a selective increase in mouth fixations among children receiving 12 months of slowness intervention program (Gepner et al., 2022). These findings support the view that deliberate deceleration of audiovisual stimuli can enhance the exploration of facial communicative features in autistic children. A third RCT tested the feasibility and efficacy of a gaze-contingent ET training (GCET) as both an intervention and an outcome tool for improving attention to faces (Wang et al., 2020). At post-training, children receiving GCET showed increased fixation rates and durations on faces, suggesting that adaptive gaze training can effectively counteract the natural decline in face-directed attention typically observed in autistic children. Other interventions yielded complementary evidence. An RCT showed that a 5-month classroom-based social program combined with caregiver training improved children's attention to faces during dynamic action understanding tasks relative to caregiver training alone, suggesting that structured group-based interventions can enhance social cue processing (Haworth et al., 2018). However, this conclusion is hampered by the small sample size and high rates of missing ET data. More recently, Bast et al. (2025) investigated the effect of the Frankfurt Early Intervention Programme for ASD (A-FFIP) on the development of reactive joint attention (RJA) in autistic preschoolers, using an ET paradigm combined with pupillometry to assess gaze following and arousal regulation. Results showed that A-FFIP increased RJA likelihood (proportion of response) with effects persisting for up to two years post-intervention, and cascading into improved social responsiveness, partly mediated by arousal regulation. Collectively, these trials indicate that direct interventions may shape visual attention in autism, though effects are inconsistent and context-dependent. Paradigm choice appears critical: while tasks assessing spontaneous social preference showed limited sensitivity to change, interventions that manipulated stimulus properties (Gepner et al., 2022) or embed social cues in naturalistic contexts (Haworth et al., 2018; Bast et al., 2025) seem more promising. Nevertheless, methodological constraints—particularly small effective sample sizes, heterogeneous designs, and lack of long-term follow-up limit interpretation. These studies therefore underscore both the potential of direct interventions to influence visual attentional patterns and the need for more robust, ecologically valid research to confirm whether such gains generalize beyond the laboratory.

Pharmacological interventions. Three RCTs incorporated ET outcome measures to determine the efficacy of pharmacological interventions in young children with autism. A double-blind, crossover study evaluated the effect of intranasal oxytocin administration paired with positive social interaction on social attention and autism-related symptomatology (Le et al., 2022). Intranasal oxytocin significantly increased fixation duration on social versus geometric stimuli compared to placebo, and selectively enhanced attention to the eyes of angry,

happy, and neutral faces, while reducing attention to fearful faces. These results suggest that oxytocin may modulate social attention in an emotion-specific manner, potentially enhancing social engagement of children with autism. By contrast, a phase II double-blind, placebo--controlled trial testing the efficacy of umbilical cord blood (CB) infusions in 180 preschool children with autism reported no overall treatment effects on ET-derived social attention during a dyadic bid task (Dawson et al., 2020). Unexpectedly, participants spent more time looking at toys during the dyadic bid after intervention. Subgroup analyses suggested improved sustained attention to complex social stimuli in CB-treated children without intellectual disability. Finally, a third trial investigated the effects of sertraline, a selective serotonin reuptake inhibitor, on language development and social behavior in 58 young children with autism (Potter et al., 2019). Using a passive-viewing eye-tracking (PVET) paradigm designed to measure receptive vocabulary processing, fixation duration on the correct picture was recorded as a measure of word recognition ability. The study found no intervention effects on fixation patterns, although usable ET data were available for only 14 (7 vs. 7) children, severely limiting interpretability. These results were consistent with previous evidence that sertraline did not significantly improve receptive language or social behavior (Williams et al., 2010). These trials highlight that no pharmacological approach can yet be considered broadly effective. Oxytocin shows the most consistent promise, though with nuanced effects. In contrast, umbilical CB infusion yielded null findings aside from exploratory subgroup effects, while sertraline showed no evidence of improvement in receptive language or social behavior, further constrained by very limited analyzable data. Overall, these findings suggest that pharmacological approaches may differentially influence ET outcomes depending on their proximity to social-attentional mechanisms. Methodological constraints—including small effective sample sizes, reliance on diverse ET paradigms, and exploratory subgroup analyses—limit firm conclusions and underscore as well the need to better identify individual response profiles before the interventions can be widely recommended.

3.4. *Exploratory meta-analysis and certainty of evidence*

An exploratory meta-analysis was conducted for the subset of studies that reported sufficiently comparable ET outcomes (Fig. 3). Although based on a small number of studies, the pooled estimate of two RCTs (n = 101) showed a moderate intervention effect (SMD = 0.42, 95 % CI [0.02, 0.82]) on attention disengagement, suggesting that early interventions might enhance the ability to shift attention away and increase attentional flexibility, enabling more efficient allocation of visual attention. Evidence for fixation-based measures was less consistent. For total fixation time (TFT) to the eyes, pooled results from two RCTs

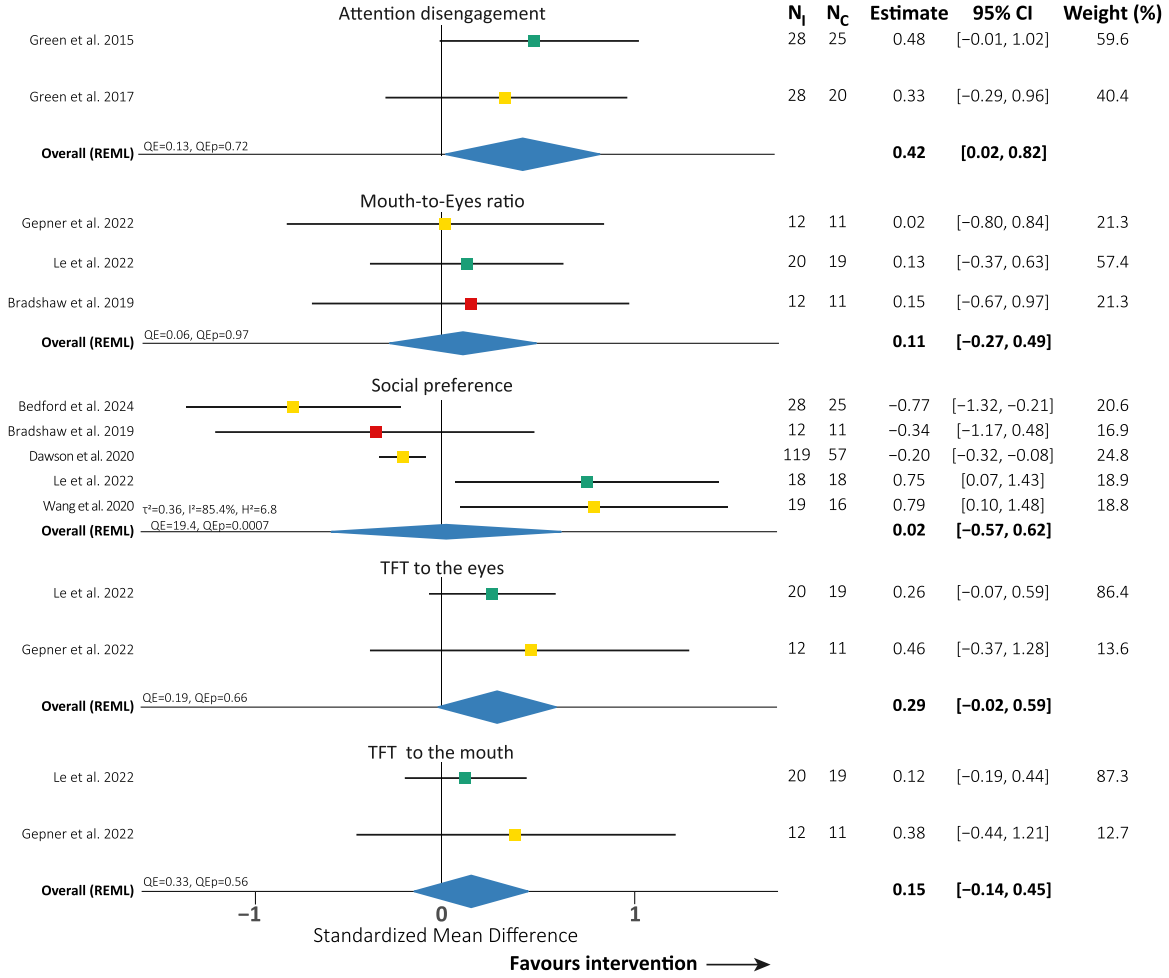

Fig. 3. Exploratory meta-analysis of early intervention effects on ET outcome measures in young children with or at risk of autism. Forest plots display standardized mean differences (SMD) with 95 % confidence intervals (CI) for five ET outcomes: attention disengagement, mouth-to-eyes ratio, social preference, total fixation time (TFT) to the eyes and to the mouth. Pooled estimates were calculated using random-effects models with restricted maximum likelihood (REML). The blue diamond represents the overall summary effect size for each outcome. Individual study effect sizes are shown as squares, and their color denotes risk of bias (green = low, yellow = moderate, red = high). Heterogeneity statistics (τ^2, I^2, H^2, and Cochran's Q with p-value) are reported where applicable. Weight of each study was computed per ET outcome using $1/(vi+\tau^2)$, in which vi is the sampling variance of the effect size. Positive SMD values indicate a beneficial effect of intervention.

(n = 62) suggested a small positive effect (SMD = 0.29, 95 % CI [-0.02, 0.59]), whereas TFT to the mouth (two RCTs, n = 62) and mouth-to-eyes ratio (three RCTs, n = 85) showed no differences between intervention and control groups (SMD = 0.15, 95 % CI [-0.14, 0.45] and SMD = 0.11, 95 % CI [-0.27, 0.49], respectively). For social preference, based on five RCTs (n = 323), the pooled estimate indicated no consistent effect (SMD = 0.02, 95 % CI [-0.57, 0.62]), with substantial heterogeneity (I^2 = 85.4 %). Given the heterogeneity of intervention strategies and the low number of studies, these results should be interpreted with caution but provide a preliminary quantitative complement to the narrative synthesis.

The GRADE assessment (Table 2) showed that certainty of the evidence was consistently limited. Certainty of evidence for attentional disengagement was rated low due to imprecision and limited number of trials. All other outcomes were rated very low certainty, reflecting methodological shortcomings, inconsistency across studies, and wide confidence intervals that spanned both null and clinically meaningful effects. These ratings indicate that further research is very likely to influence current effect estimates.

3.5. Registered RCTs

We identified 24 additional registered RCTs evaluating early intervention programs in young children with autism that mentioned the use of an ET outcome measure (Table 3). Of these, 11 RCTs showed a status of *completed/terminated,* in which ET assessments were undertaken yet not reported in the results, remaining thus unpublished. The status of two of these studies is "unknown", and another RCT (NCT05413187) was discontinued due to the Covid-19 pandemic. The authors of one of the completed RCTs published a protocol for integrated standardized autism screening, ESDM intervention and long-term outcomes (McClure et al., 2021). The remaining thirteen registered RCTs are labelled as "ongoing", and two of them have their respective protocol published: the first protocol article outlines a study evaluating the efficacy of ENACT (ENvironmental enrichment for infants parenting with Acceptance and Commitment Therapy) (Whittingham et al., 2020), while the second published protocol describes an ongoing RCT conducted by our research team to evaluate the efficacy of E-Coaching, a remotely delivered PMI (Peter et al., 2025). The targeted population in the registered RCTs were infants at high risk of autism in 25 % (6/24) of the studies, whereas older children with autism were recruited in the remaining 18 RCTs. Six studies assessed the efficacy of PMIs, six RCTs evaluated therapist-delivered interventions, and four additional studies examined the efficacy of pharmacological interventions. Six RCTs involved combined intervention approaches (e.g., direct plus PMI, or PMI plus oxytocin). Finally, two personalized intervention strategies were explored in the two remaining RCTs, evaluating the effects of precise neuromodulation (ChiCTR-INR-17012854) and neurostimulation (ChiCTR2200061683) procedures. With respect to the ET outcome, 42 % (10/24) of registered RCTs used social attention tasks, whilst 17 % (4/24) employed nonsocial attention disengagement, and 4 % (1/24) measured visual preferences. The remaining RCTs employed other ET paradigms (e.g., predictive gaze) or no details were provided.

4. Discussion

The aim of this systematic review was to identify, appraise and summarize the existing empirical evidence on the use of ET as an outcome measure in RCTs assessing the efficacy of early pharmacological and psychosocial interventions in young children with autism. We identified eleven reports all published after 2015, highlighting the timeliness of this review. We found an overall low-to-moderate risk of bias in most of the selected studies using the Cochrane RoB-2 tool. The assessment of visual attentional patterns encompassed a wide range of ET systems and paradigms. Collectively, a part of the literature reviewed here provides reliable indices of enhanced social and nonsocial attention, ultimately associated with broader developmental gains. While our systematic search also revealed notable research gaps, ET proved to be a clinically relevant and change-sensitive tool to demonstrate the effectiveness of early intervention programs in several contexts. Exploratory findings from the meta-analysis suggest a potential intervention benefit for attention disengagement, whereas results for other outcomes remain inconclusive. The overall certainty of evidence was rated low to very low, primarily due to the small number of available studies, which limited precision and increased susceptibility to bias and inconsistency. Despite these limitations, these findings underscore ET's potential as a

Table 2
GRADE Summary of findings for ET outcomes. Patient or population: Young children at risk of or diagnosed with autism. Intervention: early intervention programs. Comparison: treatment as usual (TAU)/placebo.

Outcome	Anticipated absolute effects* SMD [95 % CI]		Number of participants (studies)	τ^2/ I^2(%)/ H^2	RoB (Low/ Mod/High)	Certainty of Evidence	Comments
	Risk for TAU	Risk for intervention					
Attention disengagement	-	0.42 [0.02, 0.82]	101 (2 RCTs)	0/0/1	1/1/0	⊕⊕○○ Low[b,d]	May improved attention disengagement but confidence in the effect estimate is still limited: the true effect may be substantially different
Mouth-to-eyes ratio	-	0.11 [-0.27, 0.49]	85 (3 RCTs)	0/0/1	1/1/1	⊕○○○ Very low [a,b,e]	Very little confidence in the effect estimate due to risk of bias, inconsistency and lack of precision
Social preference	-	0.02 [-0.57, 0.62]	323 (5 RCTs)	0.36/ 85.4/6.8	1/3/1	⊕○○○ Very low [a,c,e]	Very little confidence in the effect estimate due to risk of bias, serious inconsistency and lack of precision
TFT to the eyes	-	0.29 [-0.02, 0.59]	62 (2 RCTs)	0/0/1	1/1/0	⊕○○○ Very low [b,d,e]	There may be little or no difference due to inconsistency and lack of precision
TFT to the mouth	-	0.15 [−0.14, 0.45]	62 (2 RCTs)	0/0/1	1/1/0	⊕○○○ Very low [b,d,e]	Very little confidence in the effect estimate due to inconsistency and lack of precision

Abbreviations: SMD= Standardized mean difference, CI= Confidence Interval, τ2= estimated variance of the true effect size across studies, I^2= % of total variation that is due to heterogeneity, H^2=Between-study heterogeneity, RoB= Risk of Bias, TAU= treatment as usual, TFT= total fixation time.
Note: *The risk in the intervention group (and its 95 % confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention.
[a] High Risk of Bias present
[b] Inconsistency (k ≤ 3; heterogeneity not reliably assessable)
[c] Serious inconsistency for I^2 ≥75 % or τ ≥0.30
[d] Imprecision (k ≤ 2)
[e] Imprecision (95 % CI spans values including no effect and clinically meaningful effects)

Table 3
Registered completed and ongoing studies.

Author, Year (start)	Code	Population	Age	Intervention & Comparator	ET Outcome Measure	Current status (date)
Completed or discontinued RCTs						
Hardan AY, et al. (2013)	NCT02037022	ASD, N = 48	2–5 years	PRT ***vs*** No intervention	Social attention; word-learning	Completed (2017)
Caetano S, et al. (2014)	NCT02235467	ASD, N = 67	3–7 years	ABA parent training ***vs*** No Intervention	Latency, duration, and sequence of visual fixations	Completed (2014)
Frazier TW., et al. (2014)	NCT02202421	ASD, N = 21	24–71 months	T-ABA Parent Group + Individual Therapy ***vs*** T-ABA Parent Training Only	Social attention	Terminated (2017)
Gengoux G, et al. (2017)	NCT03177525	ASD, N = 22	4–6 years	Social SUCCESS ***vs*** Waiting list	Social attention, not specified	Completed (2021)
Brian J, et al. (2017)	NCT03215394	Suspected ASD, N = 150	12–30 months	Enhanced Social ABCs ***vs*** Standard Social ABCs ***vs*** TAU	Attention disengagement	Completed (2020)
Geoffray, MM, et al. (2017)	NCT03254264	NI	15–36 months	ESDM ***vs*** TAU	Social preference, attention disengagement, time on predictable stimuli	Terminated (2022)
Tager-Flusberg HB, et al. (2018)	NCT03427138	ASD/suspected ASD, N = 100,	18–59 months	JASPER ***vs*** Parent Education	Predictive gaze	Unknown
Lense M, et al. (2018)	NCT03560297	With or without ASD, N = 70	20–72 months	SeRenade Program ***vs*** Waiting list	Social attention	Completed (2020)
Bar-Lev Schleider, L, et al. (2019)	NCT05413187	N = 0	2–8 years	Medical Grade Cannabis oil ***vs*** Placebo	Social attention	Withdrawn
Sijia Gu, et al. (2020)	NCT04319640	ASD, N = 160	24–59 months	CBT for insomnia ***vs*** Psychoeducation	Synchronized ET	Unknown
McClure et al. (2021)	https://doi.org/10.1186/s13063–021–05286–6	Children, N = 2087	16–58 months	Enhanced early detection +ESDM ***vs*** TAU	Social orienting, social cognition, social motivation	Completed (2024) **Published protocol**
Ongoing RCTs						
Gustella A, et al. (2018)	ACTRN12618001029280	ASD, N = 132	3–5 years	P-ESDM + oxytocin ***vs*** P-ESDM	Social attention	Recruiting
Qiu J, et al. (2017)	ChiCTR-INR-17012854	ASD, N = 40	2–6 years	TDCS ***vs*** Pseudo TDCS	NI	Recruiting
Whittingham K, et al. (2020)	https://doi.org/10.1136/bmjopen−2019–034315	High risk of ASD, N = 66	0–12 months	ENACT ***vs*** TAU	Attention disengagement	Ongoing **Published protocol**
Rongfang Hu, et al. (2020)	ChiCTR2000028896	ASD, N = 140	3–7 years	Sandplay ***vs*** Creative Handicraft and Painting ***vs*** Dance Movement Therapy ***vs*** Control	NI	Recruiting
Duan X, et al. (2022)	ChiCTR2200061683	ASD, N = 100	2–8 years	Personalized precision stimulate target located by MRI data ***vs*** Traditional Target Group	NI	Recruiting
Sun J, et al. (2022)	ChiCTR2200066545	ASD, N = 120	2–7 years	Oral of Lactobacillus plantarum PS128 ***vs*** Placebo	Arrival time of the first gaze; duration of the first fixation point	Recruiting
Roeyers H, et al. (2022)	NCT05208411	High risk of ASD, N = 80	9–18 months	ImPACT ***vs*** No intervention	Attention disengagement	Recruiting
Sun J., et al. (2023)	ChiCTR2300070783	ASD, N = 90	2–6 years	High dose L. reuteri CICC6132 ***vs*** Low-dose vs rehabilitation training	NI	Recruiting
Liu G., et al. (2023)	ChiCTR2300068699	ASD, N = 100	3–7 years	Expressive Art Therapy ***vs*** Control	NI	Not yet recruiting
Lense M., et al. (2023)	NCT05880225	ASD, N = 40	18–36 months	meRIT ***vs*** RIT	Trajectories of rhythmically entrained eye-looking to predictable and unpredictable child-directed singing	Recruiting
Kendrick K., et al. (2024)	ChiCTR2400090166	ASD, N = 180	2–6 years	Oxytocin low dose ***vs*** medium dose ***vs*** high dose ***vs*** placebo	Social attention	Recruiting

(*continued on next page*)

Table 3 (*continued*)

Author, Year (start)	Code	Population	Age	Intervention & Comparator	ET Outcome Measure	Current status (date)
Rollins P., et al. (2024)	NCT06596226	Suspected ASD, N = 80	16–30 months	Pathways with Mutual Gaze Protocol ***vs*** Pathways without Mutual Gaze Protocol	Social attention	Recruiting
Peter C., et al. (2025)	https://doi.org/10.1111/nyas.15320	ASD, N = 99	24–48 months	E-Coaching ***vs*** PACT ***vs*** TAU	Fixation on face vs toys, shared attention (head-mounted ET)	Recruiting **Published protocol**

NI= no information, ASD= Autism Spectrum Disorder, PRT=Pivotal Response Training, ABA= Applied Behavior Analysis, ESDM = Early Start Denver Model, TAU= Treatment as usual, Jasper = Joint Attention, Symbolic Play and Engagement Regulation, ENACT= ENvironmental enrichment for infants parenting with Acceptance and Commitment Therapy, CBT= Cognitive Behavioral Intervention, TDCS =Transcranial direct current stimulation, meRIT= Music-Enhanced Reciprocal Imitation Training, RIT=Reciprocal Imitation Training PACT= Peadiatric Autism Communication Therapy

valuable outcome measure and underscore the need for larger, methodologically rigorous trials with harmonized ET paradigms.

ET is a blindly assessed objective measure that provides millisecond-level insights into how visual attentional patterns fluctuate. The use of ET in infant populations has demonstrated its accessibility and feasible implementation even with non-verbal populations and/or those with limited compliance (Perkovich et al., 2024; Tager-Flusberg et al., 2017). A suitable ET outcome requires as well to be sensitive to fine changes over time and able to detect clinically meaningful differences, which has proven difficult to achieve. Our review provides preliminary support for the potential of ET to meet several of the above criteria. In particular, the increase of attention disengagement in early childhood represents a reliable early indicator for a subsequent diagnosis of autism (Hou et al., 2024b; Sacrey et al., 2013). As such, the subtle yet solid enhancement of nonsocial attentional flexibility following iBASIS-VIPP intervention (Green et al., 2015) stands out for its potential cross-domain impact on social communication competences (Elsabbagh et al., 2011). Changes in attentional flexibility following the iBASIS-VIPP might thus enable more efficient allocation of visual attention, a pivotal skill for social interactions and the development of joint attention. Nevertheless, demonstrating sensitivity to change over longer time spans to validate sustained intervention effects has proven more challenging. In this respect, Bast et al. (2025) reports consolidated, persisting changes in joint attention beyond the immediate intervention period. The durability of these effects—extending up to two years post-intervention—distinguishes this study from other trials, which have rather shown short-term or context-specific improvements. This suggests that sustained modification of attentional processes in autism may be achievable, though evidence remains preliminary and requires replication. In pharmacological interventions, increased visual attention to dynamic social stimuli and to the eyes of emotional faces following oxytocin administration was further supported by a significant reduction in social communication deficits (Le et al., 2022). Globally, the three abovementioned examples stem from solid studies with a low risk of methodological bias, confirming the ability of ET to capture clinically meaningful intervention effects and thereby strengthening its clinical pertinence. ET was also sensitive enough to ascertain changes associated with nonpharmacological intervention strategies delivered directly to the child, such as increased fixation on faces following a gaze-contingent training (Wang et al., 2020), and longer times looking at the mouth after a slowness intervention program (Gepner et al., 2022).

In other studies, however, ET outcomes showed unexpected results against the pre-specified hypotheses, such as the reduced dwell time after iBASIS-VIPP intervention (Bedford et al., 2024), indicating that ET could still potentially detect changes that are not specifically driven by the effect of the intervention. Importantly, several of the selected studies with inconsistent or unexpected results presented methodological shortcomings, such as bias arising from the randomization process, selective reporting of outcomes, deviations from intended interventions and missing outcome data. To illustrate, three studies targeting social preference presented counterintuitive intervention effects (Bedford et al., 2024; Bradshaw et al., 2019; Dawson et al., 2020). All exhibited deficiencies in the randomization process, and two were further limited by selective outcome reporting. Such risks of bias not only compromise internal validity but also undermine confidence in whether observed effects reflect genuine intervention impact or methodological artifacts (Chandler et al., 2019). Within the context of this review, these limitations might artificially inflate or mask associations, leading to inconsistent conclusions about ET's reliability and sensitivity as an outcome measure. Moreover, the nature of the intervention approaches under study can further challenge the interpretation of the effects. Some interventions, such as umbilical CB infusion, are presumed to be biologically distant from directly influencing social attentional processes, which are the primary constructs typically indexed by ET. By contrast, interventions more closely tied to social functioning—such as oxytocin administration or PMIs—are theoretically and empirically more likely to elicit measurable changes in ET-derived indices of social attention (Green et al., 2015; Le et al., 2022; Yamasue et al., 2020). This divergence highlights the need to carefully match outcome measures with the putative mechanisms targeted by the intervention.

While some of the abovementioned ET paradigms seem reliable enough to effectively measure changes over time, other studies also demonstrated independent replication of the potential utility of social preference indices as diagnostic biomarkers for autism (Keehn et al., 2024; Wen et al., 2022). Consistently, the two RCTs that included a cohort of TD children provided support for the sensitivity of ET to detect between-group differences in social attention when compared to normative developmental trajectories (Bradshaw et al., 2019; Wang et al., 2020). Alternatively, certain ET measures proved to be more effective at predicting the developmental course of clinical outcomes, such as eye-to-mouth ratio (Bradshaw et al., 2019; Young et al., 2009). In this sense, there is growing interest in the search for baseline predictors of optimal developmental progress following early intervention (Asta and Persico, 2022), with a few studies using ET for this purpose with contrasting results. For example, the assessment of learning profiles associated with response to ESDM in group settings showed that more attention to faces was not associated with greater developmental gains (Vivanti et al., 2013). On the other hand, other studies revealed that higher levels of attention to faces were predictive of developmental progress following ESDM intervention (Latrèche et al., 2021). Finally, the use of ET to develop a stratification biomarker is another promising avenue with the potential to challenge the wild heterogeneity of autism-related manifestations. Clustering individuals within RCTs into subgroups relevant to predicting or evaluating interventions has already yielded encouraging results, such as the enhanced social attention following CB treatment in a subgroup of autistic children without intellectual disability (Dawson et al., 2020). At present, only three biomarkers for neuropsychiatric conditions have advanced to the Letter of Intent stage with the Food and Drug Administration (FDA), none of which has yet progressed to the qualification stage. One of these

(DDT-BMQ-000093) concerns an oculomotor index of attention to human faces as a stratification biomarker to reduce heterogeneity of individuals with autism participating in clinical trials (Shic et al., 2022).

Several future directions emerge from our literature assessment. The integration of ET and ML models has shown significant potential in advancing the early detection of autism (Jaradat et al., 2025; Kong et al., 2022; Meng et al., 2023; Wei et al., 2024). These technologies provide data-driven alternatives to traditional behavioral assessments, capturing even slight variations in social and nonsocial attentional patterns that may not be readily observable through clinical evaluations alone. As ML algorithms continue to evolve, they could serve a similar purpose in the use of ET for intervention monitoring, helping to refine treatment strategies by identifying individualized trajectories of developmental progress. Secondly, recent advances in mobile ET (MET) devices offer several advantages over traditional screen-based settings, by enabling the assessment of visual attention in naturalistic, dynamic environments (Aukerman et al., 2025; Fu et al., 2024). MET captures gaze patterns from a first-person perspective, allowing the study of visual attention as it integrates with motor actions and social behavior. Unlike conventional stationary systems, MET facilitates the study of how attention unfolds and interacts with the environment in real time. Technological advances have extended the applicability of MET to pediatric populations, thereby enriching developmental research. Future perspectives should focus on ensuring that both ML and MET technologies are not only accurate and scalable, but also accessible and interpretable for widespread adoption in real-world clinical settings.

In considering ET as a potential biomarker of response to intervention, practical challenges must be acknowledged. High exclusion rates and missing data in very young children remain common, reducing sample representativeness and statistical power. Another concern involves barriers to implementation in low-resource settings, given the costs of specialized equipment, requirements for infrastructure and training needs. Nevertheless, recent advances—such as the above-mentioned development of low-cost MET systems (Dalmaijer, 2014; Hessels and Hooge, 2019) and analytical approaches that better accommodate missing data (Wass et al., 2014)—are beginning to mitigate these challenges. These innovations shed light on the realistic pathways through which ET can become a broadly feasible and impactful outcome measure. Progress in standardization, data quality, and accessibility will be critical to ensure that ET can be applied effectively across diverse research and clinical contexts.

The present review has both strengths and limitations. The employment of a meticulous search strategy and the rigorous methodological framework constitute remarkable strengths of the study. Some limitations should however be noted. Publication bias is a potential concern, as numerous studies remain unpublished, which may result in an overrepresentation of positive findings while negative or unsignificant results are underreported. The inclusion criteria focused on a specific age range, leading to the exclusion of studies that might have offered valuable insights into older populations. Limited access to ongoing and unpublished studies also represents a constraint, as our analysis was exclusively based on publicly available data without directly contacting authors for additional evidence. While our exploratory meta-analysis offers an initial quantitative indication of effect sizes, the analysis was constrained by the limited number of studies and variability in outcome reporting. Future research using standardized ET metrics and more consistent reporting will be critical for enabling a more conclusive meta-analytic synthesis. Lastly, language restrictions may have led to the omission of relevant studies published in languages other than English.

Our review leaves some questions unanswered and supports that further work is warranted to achieve greater consistency in neurobehavioral biomarkers of response to early intervention in autism. The standardization of ET measures is challenging, and there are frequently elevated rates of missing data. Providing researchers with unambiguous guidelines on the strengths and limitations of prevailing ET metrics, along with a comprehensive framework to develop a standardized set of outcome measures, may ultimately ensure the generalizability of the means to detect the neurobehavioral effects of early interventions. The establishment of such a gold standard would serve to increase the replicability of results and enhance the comparability of data across studies and disciplines. As ongoing trials continue to explore ET as a tool for measuring intervention outcomes, future research should investigate how these findings translate to real-world clinical practice.

Author contributions

CP, NC and BRH conceived the review and oversaw all aspects of the work. CP and BRH conducted the screening and data extraction. CP, MPA and BRH performed the critical appraisal. JRA mentored the databases search, generated the PRISMA diagram and wrote the methods section. JAO and EA contributed to the editing of the manuscript. CP, MPA and BRH wrote the review with input from all the authors.

Declaration of Competing interest

NC receives support from *Fondation Dora* and *Fondation Hoffmann*; she declares having no other competing interests. The other authors declare no competing interests.

Acknowledgements

We are particularly thankful to the finantial support of Fondation Dora and Fondation Hoffmann-La Roche. We also thank the reviewer for her/his constructive and thoughtful comments.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.neubiorev.2025.106391.

Data availability

No data was used for the research described in the article.

References

Accordino, R.E., Kidd, C., Politte, L.C., Henry, C.A., McDougle, C.J., 2016. Psychopharmacological interventions in autism spectrum disorder. Expert Opin. Pharm. 17, 937–952. https://doi.org/10.1517/14656566.2016.1154536.

Alsharif, N., Al-Adhaileh, M.H., Al-Yaari, M., Farhah, N., Khan, Z.I., 2024. Utilizing deep learning models in an intelligent eye-tracking system for autism spectrum disorder diagnosis. Front Med. 11, 1436646. https://doi.org/10.3389/fmed.2024.1436646.

Alvares, G.A., Chen, N.T.M., Notebaert, L., Granich, J., Mitchell, C., Whitehouse, A.J.O., 2019. Brief social attention bias modification for children with autism spectrum disorder. Autism Res. 12, 527–535. https://doi.org/10.1002/aur.2067.

American Psychiatric Association, 2022. Diagnostic And Statistical Manual Of Mental Disorders, Fifth ed. American Psychiatric Association [APA]. Text Revision (DSM-5-TR).

Anagnostou, E., Jones, N., Huerta, M., Halladay, A.K., Wang, P., Scahill, L., Horrigan, J. P., Kasari, C., Lord, C., Choi, D., Sullivan, K., Dawson, G., 2015. Measuring social communication behaviors as a treatment endpoint in individuals with autism spectrum disorder. Autism 19, 622–636. https://doi.org/10.1177/1362361314542955.

Aslin, R.N., 2012. Infant eyes: a window on cognitive development. Infancy 17, 126–140. https://doi.org/10.1111/j.1532-7078.2011.00097.x.

Asmetha Jeyarani, R., Senthilkumar, R., 2023. Eye tracking biomarkers for autism spectrum disorder detection using machine learning and deep learning techniques: review. Res. Autism Spectr. Disord. 108, 102228. https://doi.org/10.1016/j.rasd.2023.102228.

Asta, L., Persico, A.M., 2022. Differential predictors of response to early start Denver model vs. Early intensive behavioral intervention in young children with autism spectrum disorder: a systematic review and Meta-Analysis. Brain Sci. 12. https://doi.org/10.3390/brainsci12111499.

Atkinson Jr, A.J., Colburn, W.A., DeGruttola, V.G., DeMets, D.L., Downing, G.J., Hoth, D. F., Oates, J.A., Peck, C.C., Schooley, R.T., 2001. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin. Pharmacol. Ther. 69, 89–95. https://doi.org/10.1067/mcp.2001.113989.

Aukerman, J.F., Newland, J., Scott, L.S., 2025. Eye tracking for infant populations: a comparison of methods and best practices: eye tracking for infant populations. ACM Trans. Appl. Percept. https://doi.org/10.1145/3719009.
Balshem, H., Helfand, M., Schunemann, H.J., Oxman, A.D., Kunz, R., Brozek, J., Vist, G. E., Falck-Ytter, Y., Meerpohl, J., Norris, S., Guyatt, G.H., 2011. GRADE guidelines: 3. Rating the quality of evidence. J. Clin. Epidemiol. 64, 401–406. https://doi.org/10.1016/j.jclinepi.2010.07.015.
Bast, N., Polzer, L., Raji, N., Schnettler, L., Kleber, S., Lemler, C., Kitzerow-Cleven, J., Kim, Z., Schaer, M., Freitag, C.M., 2025. Early intervention increases reactive joint attention in autistic preschoolers with arousal regulation as mediator. Eur. Child Adolesc. Psychiatry. https://doi.org/10.1007/s00787-025-02738-1.
Bedford, R., Green, J., Gliga, T., Jones, E.H., Elsabbagh, M., Pasco, G., Wan, M.W., Slonims, V., Charman, T., Pickles, A., Johnson, M.H., Team, B., 2024. Parent-mediated intervention in infants with an elevated likelihood for autism reduces dwell time during a gaze-following task. Autism Res. 17, 2346–2354. https://doi.org/10.1002/aur.3223.
Borissov, N., Haas, Q., Minder, B., Kopp-Heim, D., von Gernler, M., Janka, H., Teodoro, D., Amini, P., 2022. Reducing systematic review burden using deduklick: a novel, automated, reliable, and explainable deduplication algorithm to foster medical research. Syst. Rev. 11, 172. https://doi.org/10.1186/s13643-022-02045-9.
Bradshaw, J., Shic, F., Holden, A.N., Horowitz, E.J., Barrett, A.C., German, T.C., Vernon, T.W., 2019. The use of eye tracking as a biomarker of treatment outcome in a pilot randomized clinical trial for young children with autism. Autism Res. 12, 779–793. https://doi.org/10.1002/aur.2093.
Brian, J.A., Smith, I.M., Zwaigenbaum, L., Bryson, S.E., 2017. Cross-site randomized control trial of the social ABCs caregiver-mediated intervention for toddlers with autism spectrum disorder. Autism Res. 10, 1700–1711. https://doi.org/10.1002/aur.1818.
Califf, R.M., 2018. Biomarker definitions and their applications. Exp. Biol. Med. 243, 213–221. https://doi.org/10.1177/1535370217750088.
Campbell, D.J., Shic, F., Macari, S., Chawarska, K., 2014. Gaze response to dyadic bids at 2 years related to outcomes at 3 years in autism spectrum disorders: a subtyping analysis. J. Autism Dev. Disord. 44, 431–442. https://doi.org/10.1007/s10803-013-1885-9.
Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V., 2019. Cochrane handbook for systematic reviews of interventions. Hoboken Wiley 4, 14651858.
Cheuk, D.K., Wong, V., Chen, W.X., 2011. Acupuncture for autism spectrum disorders (ASD). Cochrane Database Syst. Rev. https://doi.org/10.1002/14651858.CD007849.pub2.
Chita-Tegmark, M., 2016. Social attention in ASD: a review and meta-analysis of eye-tracking studies. Res. Dev. Disabil. 48, 79–93. https://doi.org/10.1016/j.ridd.2015.10.011.
Cochrane, C., 2024. Cochrane Handb. Syst. Rev. Interv. Version 6. 5 (Update August 2024).
Crank, J.E., Sandbank, M., Dunham, K., Crowley, S., Bottema-Beutel, K., Feldman, J., Woynaroski, T.G., 2021. Understanding the effects of naturalistic developmental behavioral interventions: a project AIM Meta-analysis. Autism Res. 14, 817–834. https://doi.org/10.1002/aur.2471.
Dalmaijer, E., 2014. Is the low-cost EyeTribe eye tracker any good for research? PeerJ Prepr.
Dawson, G., 2008. Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. Dev. Psychopathol. 20, 775–803. https://doi.org/10.1017/s0954579408000370.
Dawson, G., Bernier, R., Ring, R.H., 2012. Social attention: a possible early indicator of efficacy in autism clinical trials. J. Neurodev. Disord. 4, 11. https://doi.org/10.1186/1866-1955-4-11.
Dawson, G., Meltzoff, A.N., Osterling, J., Rinaldi, J., Brown, E., 1998. Children with autism fail to orient to naturally occurring social stimuli. J. Autism Dev. Disord. 28, 479–485. https://doi.org/10.1023/a:1026043926488.
Dawson, G., Sun, J.M., Baker, J., Carpenter, K., Compton, S., Deaver, M., Franz, L., Heilbron, N., Herold, B., Horrigan, J., Howard, J., Kosinski, A., Major, S., Murias, M., Page, K., Prasad, V.K., Sabatos-DeVito, M., Sanfilippo, F., Sikich, L., Simmons, R., Song, A., Vermeer, S., Waters-Pick, B., Troy, J., Kurtzberg, J., 2020. A phase II randomized clinical trial of the safety and efficacy of intravenous umbilical cord blood infusion for treatment of children with autism spectrum disorder. J. Pediatr. 222, 164–173. https://doi.org/10.1016/j.jpeds.2020.03.011.
Dawson, G., Sun, J.M., Davlantis, K.S., Murias, M., Franz, L., Troy, J., Simmons, R., Sabatos-DeVito, M., Durham, R., Kurtzberg, J., 2017. Autologous cord blood infusions are safe and feasible in young children with autism spectrum disorder: results of a Single-Center phase I Open-Label trial. Stem Cells Transl. Med. 6, 1332–1339. https://doi.org/10.1002/sctm.16-0474.
Delrobaee, A., Tehrani-Doost, M., Joghataei, M., Razjouyan, K., Benaraki, A.K., Mollajani, R., 2020. Bumetanide-related effect on attention to face parts in children and adolescents with autism spectrum disorder using the eye tracking algorithm. Adv. Cogn. Sci. 21, 84–93. https://doi.org/10.30699/icss.21.4.84.
Egger, M., Davey Smith, G., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. BMJ 315, 629–634. https://doi.org/10.1136/bmj.315.7109.629.
Elsabbagh, M., Fernandes, J., Webb, S.J., Dawson, G., Charman, T., Johnson, M.H., 2013. Disengagement of visual attention in infancy is associated with emerging autism in toddlerhood. Biol. Psychiatry 74, 189–194. https://doi.org/10.1016/j.biopsych.2012.11.030.
Elsabbagh, M., Holmboe, K., Gliga, T., Mercure, E., Hudry, K., Charman, T., Baron-Cohen, S., Bolton, P., Johnson, M.H., 2011. Social and attention factors during infancy and the later emergence of autism characteristics. Prog. Brain Res 189, 195–207. https://doi.org/10.1016/b978-0-444-53884-0.00025-7.
England, M.J., Butler, A.S., Gonzalez, M.L., 2015. Committee on developing Evidence-Based standards for psychosocial interventions for mental disorders; board on health sciences policy; institute of Medicine. Psychosocial interventions for mental and substance use disorders: a framework for establishing evidence-based standards. Psychosocial interventions for mental and substance use disorders: a framework for establishing evidence based standards. Institute of Medicine, Washington, DC.
Falck-Ytter, T., Bolte, S., Gredeback, G., 2013. Eye tracking in early autism research. J. Neurodev. Disord. 5, 28. https://doi.org/10.1186/1866-1955-5-28.
Fletcher-Watson, S., Hampton, S., 2018. The potential of eye-tracking as a sensitive measure of behavioural change in response to intervention. Sci. Rep. 8, 14715. https://doi.org/10.1038/s41598-018-32444-9.
Frank, M.C., Vul, E., Johnson, S.P., 2009. Development of infants' attention to faces during the first year. Cognition 110, 160–170. https://doi.org/10.1016/j.cognition.2008.11.010.
Frazier, T.W., Strauss, M., Klingemier, E.W., Zetzer, E.E., Hardan, A.Y., Eng, C., Youngstrom, E.A., 2017. A Meta-Analysis of gaze differences to social and nonsocial information between individuals with and without autism. J. Am. Acad. Child Adolesc. Psychiatry 56, 546–555. https://doi.org/10.1016/j.jaac.2017.05.005.
Fu, X., Franchak, J., MacNeill, L., Gunther, K., Borjon, J., Yurkovic-Harding, J., Harding, S., Bradshaw, J., Perez-Edgar, K., 2024. Implementing mobile eye tracking in psychological research: a practical guide. Behav. Res. Methods 56. https://doi.org/10.3758/s13428-024-02473-6.
Gepner, B., Charrier, A., Arciszewski, T., Tardif, C., 2022. Slowness therapy for children with autism spectrum disorder: a blind longitudinal randomized controlled study. J. Autism Dev. Disord. 52, 3102–3115. https://doi.org/10.1007/s10803-021-05183-6.
Gliga, T., Bedford, R., Charman, T., Johnson, M.H., 2015. Enhanced visual search in infancy predicts emerging autism symptoms. Curr. Biol. 25, 1727–1730. https://doi.org/10.1016/j.cub.2015.05.011.
GRADEpro, G., 2024. McMaster Univ. Evid. prime.
Green, J., Charman, T., Pickles, A., Wan, M.W., Elsabbagh, M., Slonims, V., Taylor, C., McNally, J., Booth, R., Gliga, T., Jones, E.J., Harrop, C., Bedford, R., Johnson, M.H., 2015. Parent-mediated intervention versus no intervention for infants at high risk of autism: a parallel, single-blind, randomised trial. Lancet Psychiatry 2, 133–140. https://doi.org/10.1016/S2215-0366(14)00091-1.
Green, J., Garg, S., 2018. Annual research review: the state of autism intervention science: progress, target psychological and biological mechanisms and future prospects. J. Child Psychol. Psychiatry 59, 424–443. https://doi.org/10.1111/jcpp.12892.
Green, J., Pickles, A., Pasco, G., Bedford, R., Wan, M.W., Elsabbagh, M., Slonims, V., Gliga, T., Jones, E., Cheung, C., Charman, T., Johnson, M., 2017. Randomised trial of a parent-mediated intervention for infants at high risk for autism: longitudinal outcomes to age 3 years. J. Child Psychol. Psychiatry Allied Discip. 58, 1330–1340. https://doi.org/10.1111/jcpp.12728.
Grzadzinski, R., Janvier, D., Kim, S.H., 2020. Recent developments in treatment outcome measures for young children with autism spectrum disorder (ASD). Semin Pedia Neurol. 34, 100806. https://doi.org/10.1016/j.spen.2020.100806.
Haworth, J.L., Libertus, K., Landa, R.J., 2018. Social intervention impacts action anticipation, goal extraction, and social interest in children with autism spectrum disorder. J. Educ. Dev. Psychol. 8. https://doi.org/10.5539/jedp.v8n2p95.
Henderson, J.M., Hollingworth, A., 1998. Eye movements during scene viewing: an overview. Eye Guid. Read. Scene Percept. 269–293. https://doi.org/10.1016/B978-008043361-5/50013-4.
Hessels, R.S., Hooge, I.T., 2019. Eye tracking in developmental cognitive neuroscience–The good, the bad and the ugly. Dev. Cogn. Neurosci. 40. https://doi.org/10.1016/j.dcn.2019.100710.
Hessels, R.S., Nuthmann, A., Nystrom, M., Andersson, R., Niehorster, D.C., Hooge, I.T.C., 2024. The fundamentals of eye tracking part 1: the link between theory and research question. Behav. Res. Methods 57, 16. https://doi.org/10.3758/s13428-024-02544-8.
Higgins J.P.T., Thomas J., Chandler J., Cumpston M., Li T., Page M.J., VA., W., Cochrane Handbook for Systematic Reviews of Interventions version 6.5 (updated August 2024), in: 2024, C. (Ed.), ⟨www.training.cochrane.org/handbook⟩.
Higgins, J.P., Thompson, S.G., 2002. Quantifying heterogeneity in a meta-analysis. Stat. Med. 21, 1539–1558. https://doi.org/10.1002/sim.1186.
Hirota, T., King, B.H., 2023. Autism spectrum disorder: a review. Jama 329, 157–168. https://doi.org/10.1001/jama.2022.23661.
Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., Van de Weijer, J., 2011. Eye Tracking: A Comprehensive Guide to Methods and Measures. oup Oxford.
Hou, W., Jiang, Y., Yang, Y., Zhu, L., Li, J., 2024b. Evaluating the validity of eye-tracking tasks and stimuli in detecting high-risk infants later diagnosed with autism: a meta-analysis. Clin. Psychol. Rev. 112. https://doi.org/10.1016/j.cpr.2024.102466.
Hou, W., Jiang, Y., Yang, Y., Zhu, L., Li, J., 2024a. Evaluating the validity of eye-tracking tasks and stimuli in detecting high-risk infants later diagnosed with autism: a meta-analysis. Clin. Psychol. Rev. 112, 102466. https://doi.org/10.1016/j.cpr.2024.102466.
Hustyi, K.M., Ryan, A.H., Hall, S.S., 2023. A scoping review of behavioral interventions for promoting social gaze in individuals with autism spectrum disorder and other developmental disabilities. Res. Autism Spectr. Disord. 100. https://doi.org/10.1016/j.rasd.2022.102074.
Jaradat, A.S., Wedyan, M., Alomari, S., Barhoush, M.M., 2025. Using machine learning to diagnose autism based on eye tracking technology. Diagnostics 15, 66. https://doi.org/10.3390/diagnostics15010066.
Javitt, D.C., McPartland, J.C., 2024. Neurophysiologic Biomarkers in Neuropsychiatric Disorders: Etiologic and Treatment Considerations. Springer.

Jensen, A.R., Lane, A.L., Werner, B.A., McLees, S.E., Fletcher, T.S., Frye, R.E., 2022. Modern biomarkers for autism spectrum disorder: future directions. Mol. Diagn. Ther. 26, 483–495. https://doi.org/10.1007/s40291-022-00600-7.

Jones, W., Carr, K., Klin, A., 2008. 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 65, 946–954. https://doi.org/10.1001/archpsyc.65.8.946.

Jones, W., Klin, A., 2013. Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. Nature 504, 427–431. https://doi.org/10.1038/nature12715.

Jongerius, C., Hessels, R.S., Romijn, J.A., Smets, E.M.A., Hillen, M.A., 2020. The measurement of eye contact in human interactions: a scoping review. J. Nonverbal Behav. 44, 363–389. https://doi.org/10.1007/s10919-020-00333-3.

Just, M.A., Carpenter, P.A., 1980. A theory of Reading: from eye fixations to comprehension. Psychol. Rev. 87, 329. https://doi.org/10.1037/0033-295X.87.4.329.

Kanhirakadavath, M.R., Chandran, M.S.M., 2022. Investigation of Eye-Tracking scan path as a biomarker for autism screening using machine learning algorithms. Diagnostics 12. https://doi.org/10.3390/diagnostics12020518.

Kasari, C., 2002. Assessing change in early intervention programs for children with autism. J. Autism Dev. Disord. 32, 447–461. https://doi.org/10.1023/a:1020546006971.

Keehn, B., Monahan, P., Enneking, B., Ryan, T., Swigonski, N., McNally Keehn, R., 2024. Eye-Tracking biomarkers and autism diagnosis in primary care. JAMA Netw. Open 7, e2411190. https://doi.org/10.1001/jamanetworkopen.2024.11190.

Klin, A., Jones, W., Schultz, R., Volkmar, F., Cohen, D., 2002. Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. Arch. Gen. Psychiatry 59, 809–816. https://doi.org/10.1001/archpsyc.59.9.809.

Kong, X.J., Wei, Z., Sun, B., Tu, Y., Huang, Y., Cheng, M., Yu, S., Wilson, G., Park, J., Feng, Z., Vangel, M., Kong, J., Wan, G., 2022. Different eye tracking patterns in autism spectrum disorder in toddler and preschool children. Front Psychiatry 13, 899521. https://doi.org/10.3389/fpsyt.2022.899521.

Latrèche, K., Kojovic, N., Franchini, M., Schaer, M., 2021. Attention to face as a predictor of developmental change and treatment outcome in young children with autism spectrum disorder. Biomedicines 9, 942. https://doi.org/10.3390/biomedicines9080942.

Le, J., Zhang, L., Zhao, W., Zhu, S., Lan, C., Kou, J., Zhang, Q., Zhang, Y., Li, Q., Chen, Z., Fu, M., Montag, C., Zhang, R., Yang, W., Becker, B., Kendrick, K.M., 2022. Infrequent intranasal oxytocin followed by positive social interaction improves symptoms in autistic children: a pilot randomized clinical trial. Psychother. Psychosom. 91, 335–347. https://doi.org/10.1159/000524543.

Liversedge, S.P., Findlay, J.M., 2000. Saccadic eye movements and cognition. Trends Cogn. Sci. 4, 6–14. https://doi.org/10.1016/S1364-6613(99)01418-7.

Lord, C., Charman, T., Havdahl, A., Carbone, P., Anagnostou, E., Boyd, B., Carr, T., de Vries, P.J., Dissanayake, C., Divan, G., Freitag, C.M., Gotelli, M.M., Kasari, C., Knapp, M., Mundy, P., Plank, A., Scahill, L., Servili, C., Shattuck, P., Simonoff, E., Singer, A.T., Slonims, V., Wang, P.P., Ysrraelit, M.C., Jellett, R., Pickles, A., Cusack, J., Howlin, P., Szatmari, P., Holbrook, A., Toolan, C., McCauley, J.B., 2022. The lancet commission on the future of care and clinical research in autism. Lancet 399, 271–334. https://doi.org/10.1016/S0140-6736(21)01541-5.

Lord, C., Wagner, A., Rogers, S., Szatmari, P., Aman, M., Charman, T., Dawson, G., Durand, V.M., Grossman, L., Guthrie, D., Harris, S., Kasari, C., Marcus, L., Murphy, S., Odom, S., Pickles, A., Scahill, L., Shaw, E., Siegel, B., Sigman, M., Stone, W., Smith, T., Yoder, P., 2005. Challenges in evaluating psychosocial interventions for autistic spectrum disorders. Springer 695–708. https://doi.org/10.1007/s10803-005-0017-6.

McClure, L.A., Lee, N.L., Sand, K., Vivanti, G., Fein, D., Stahmer, A., Robins, D.L., 2021. Connecting the dots: a cluster-randomized clinical trial integrating standardized autism spectrum disorders screening, high-quality treatment, and long-term outcomes. Trials 22, 319. https://doi.org/10.1186/s13063-021-05286-6.

McGlade, A., Whittingham, K., Barfoot, J., Taylor, L., Boyd, R.N., 2023. Efficacy of very early interventions on neurodevelopmental outcomes for infants and toddlers at increased likelihood of or diagnosed with autism: a systematic review and meta-analysis. Autism Res. 16, 1145–1160. https://doi.org/10.1002/aur.2924.

McPartland, J.C., 2016. Considerations in biomarker development for neurodevelopmental disorders. Curr. Opin. Neurol. 29, 118–122. https://doi.org/10.1097/wco.0000000000000300.

McPartland, J.C., Bernier, R.A., Jeste, S.S., Dawson, G., Nelson, C.A., Chawarska, K., Earl, R., Faja, S., Johnson, S.P., Sikich, L., Brandt, C.A., Dziura, J.D., Rozenblit, L., Hellemann, G., Levin, A.R., Murias, M., Naples, A.J., Platt, M.L., Sabatos-DeVito, M., Shic, F., Senturk, D., Sugar, C.A., Webb, S.J., Autism Biomarkers Consortium for Clinical, T, 2020. The autism biomarkers consortium for clinical trials (ABC-CT): scientific context, study design, and progress toward biomarker qualification. Front Integr. Neurosci. 14, 16. https://doi.org/10.3389/fnint.2020.00016.

Meng, F., Li, F., Wu, S., Yang, T., Xiao, Z., Zhang, Y., Liu, Z., Lu, J., Luo, X., 2023. Machine learning-based early diagnosis of autism according to eye movements of real and artificial faces scanning. Front Neurosci. 17, 1170951. https://doi.org/10.3389/fnins.2023.1170951.

Murias, M., Major, S., Davlantis, K., Franz, L., Harris, A., Rardin, B., Sabatos-DeVito, M., Dawson, G., 2018. Validation of eye-tracking measures of social attention as a potential biomarker for autism clinical trials. Autism Res. 11, 166–174. https://doi.org/10.1002/aur.1894.

NICE, 2021. Autism Spectr. Disord. 19s Support Manag. 2021 Surveill. Autism (NICE Guidel. CG128 CG142 CG170).

Ouzzani, M., Hammady, H., Fedorowicz, Z., Elmagarmid, A., 2016. Rayyan—a web and mobile app for systematic reviews. Syst. Rev. 5, 1–10.

Ozonoff, S., Young, G.S., Landa, R.J., Brian, J., Bryson, S., Charman, T., Chawarska, K., Macari, S.L., Messinger, D., Stone, W.L., Zwaigenbaum, L., Iosif, A.M., 2015. Diagnostic stability in young children at risk for autism spectrum disorder: a baby siblings research consortium study. J. Child Psychol. Psychiatry 56, 988–998. https://doi.org/10.1111/jcpp.12421.

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 366. https://doi.org/10.1136/bmj.n71.

Papagiannopoulou, E.A., Chitty, K.M., Hermens, D.F., Hickie, I.B., Lagopoulos, J., 2014. A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders. Soc. Neurosci. 9, 610–632. https://doi.org/10.1080/17470919.2014.934966.

Parellada, M., Andreu-Bernabeu, Á., Burdeus, M., Cáceres, A.S.J., Urbiola, E., Carpenter, L.L., Kraguljac, N.V., McDonald, W.M., Nemeroff, C.B., Rodriguez, C.I., Widge, A.S., State, M.W., Sanders, S.J., 2023. In search of biomarkers to guide interventions in autism spectrum disorder: a systematic review. Am. J. Psychiatry 180, 23–40. https://doi.org/10.1176/APPI.AJP.21100992.

van Pelt, B.J., Idris, S., Jagersma, G., Duvekot, J., Maras, A., van der Ende, J., van Haren, N.E.M., Greaves-Lord, K., 2020. The ACCEPT-study: design of an RCT with an active treatment control condition to study the effectiveness of the Dutch version of PEERS(R) for adolescents with autism spectrum disorder. BMC Psychiatry 20, 274. https://doi.org/10.1186/s12888-020-02650-9.

Perkovich, E., Laakman, A., Mire, S., Yoshida, H., 2024. Conducting head-mounted eye-tracking research with young children with autism and children with increased likelihood of later autism diagnosis. J. Neurodev. Disord. 16, 7. https://doi.org/10.1186/s11689-024-09524-1.

Peter, C., Antonietti, E., Antoniou, M.-P., Bucaille, E., Osório, J.A., Manificat, S., Rodríguez-Herreros, B., Chabane, N., 2025. E-coaching for parents of children with autism spectrum disorder: protocol for a randomized controlled trial. Ann. New Y. Acad. Sci. 1547, 143–153. https://doi.org/10.1111/nyas.15320.

Peters, J.L., Crewther, S.G., Murphy, M.J., Bavin, E.L., 2021. Action video game training improves text Reading accuracy, rate and comprehension in children with dyslexia: a randomized controlled trial. Sci. Rep. 11, 18584. https://doi.org/10.1038/s41598-021-98146-x.

Pierce, K., Conant, D., Hazin, R., Stoner, R., Desmond, J., 2011. Preference for geometric patterns early in life as a risk factor for autism. Arch. Gen. Psychiatry 68, 101–109. https://doi.org/10.1001/archgenpsychiatry.2010.113.

Pierce, K., Marinero, S., Hazin, R., McKenna, B., Barnes, C.C., Malige, A., 2016. Eye tracking reveals abnormal visual preference for geometric images as an early biomarker of an autism spectrum disorder subtype associated with increased symptom severity. Biol. Psychiatry 79, 657–666. https://doi.org/10.1016/j.biopsych.2015.03.032.

Posner, M.I., 1980. Orienting of attention. Q J. Exp. Psychol. 32, 3–25. https://doi.org/10.1080/00335558008248231.

Potter, L.A., Scholze, D.A., Biag, H.M.B., Schneider, A., Chen, Y., Nguyen, D.V., Rajaratnam, A., Rivera, S.M., Dwyer, P.S., Tassone, F., Al Olaby, R.R., Choudhary, N. S., Salcedo-Arellano, M.J., Hagerman, R.J., 2019. A randomized controlled trial of sertraline in young children with autism spectrum disorder. Front Psychiatry 10, 810. https://doi.org/10.3389/fpsyt.2019.00810.

Provenzani, U., Fusar-Poli, L., Brondino, N., Damiani, S., Vercesi, M., Meyer, N., Rocchetti, M., Politi, P., 2020. What are we targeting when we treat autism spectrum disorder? A systematic review of 406 clinical trials. Autism 24, 274–284. https://doi.org/10.1177/1362361319854641.

Rahal, R.-M., Fiedler, S., 2019. Understanding cognitive and affective mechanisms in social psychology through eye-tracking. J. Exp. Soc. Psychol. 85. https://doi.org/10.1016/j.jesp.2019.103842.

Rayner, K., 1998. Eye movements in Reading and information processing: 20 years of research. Psychol. Bull. 124, 372–422. https://doi.org/10.1037/0033-2909.124.3.372.

Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., Cook Jr., E.H., Leventhal, B.L., Pickles, A., 2006. Combining information from multiple sources in the diagnosis of autism spectrum disorders. J. Am. Acad. Child Adolesc. Psychiatry 45, 1094–1103. https://doi.org/10.1097/01.chi.0000227880.42780.0e.

Sacrey, L.A., Bryson, S.E., Zwaigenbaum, L., 2013. Prospective examination of visual attention during play in infants at high-risk for autism spectrum disorder: a longitudinal study from 6 to 36 months of age. Behav. Brain Res. 256, 441–450. https://doi.org/10.1016/j.bbr.2013.08.028.

Salvucci, D.D., Goldberg, J.H., 2000. Identifying fixations and saccades in eye-tracking protocols. Proc. 2000 Symp Eye Track. Res. Appl. 71–78. https://doi.org/10.1145/355017.355028.

Sandbank, M., Bottema-Beutel, K., Crowley, S., Cassidy, M., Dunham, K., Feldman, J.I., Crank, J., Albarran, S.A., Raj, S., Mahbub, P., Woynaroski, T.G., 2020. Project AIM: autism intervention meta-analysis for studies of young children. Psychol. Bull. 146, 1–29. https://doi.org/10.1037/bul0000215.

Sharma, R., Gulati, S., Kaur, A., Sinhababu, A., Chakravarty, R., 2022. Research discovery and visualization using ResearchRabbit: a use case of AI in libraries. COLLNET J. Scientometr. Inf. Manag. 16, 215–237. https://doi.org/10.1080/09737766.2022.2106167.

Shic, F., 2016. Eye tracking as a behavioral biomarker for psychiatric conditions: the road ahead. J. Am. Acad. Child Adolesc. Psychiatry 55, 267–268. https://doi.org/10.1016/j.jaac.2016.02.002.

Shic, F., Barney, E.C., Naples, A.J., Dommer, K.J., Chang, S.A., Li, B., McAllister, T., Atyabi, A., Wang, Q., Bernier, R., Dawson, G., Dziura, J., Faja, S., Jeste, S.S., Murias, M., Johnson, S.P., Sabatos-DeVito, M., Helleman, G., Senturk, D., Sugar, C. A., Webb, S.J., McPartland, J.C., Chawarska, K., Autism Biomarkers Consortium for Clinical, T, 2023. The selective social attention task in children with autism spectrum disorder: results from the autism biomarkers consortium for clinical trials (ABC-CT) feasibility study. Autism Res. 16, 2150–2159. https://doi.org/10.1002/aur.3026.

Shic, F., Naples, A.J., Barney, E.C., Chang, S.A., Li, B., McAllister, T., Kim, M., Dommer, K.J., Hasselmo, S., Atyabi, A., Wang, Q., Helleman, G., Levin, A.R., Seow, H., Bernier, R., Charwaska, K., Dawson, G., Dziura, J., Faja, S., Jeste, S.S., Johnson, S.P., Murias, M., Nelson, C.A., Sabatos-DeVito, M., Senturk, D., Sugar, C.A., Webb, S.J., McPartland, J.C., 2022. The autism biomarkers consortium for clinical trials: evaluation of a battery of candidate eye-tracking biomarkers for use in autism clinical trials. Mol. Autism 13, 15. https://doi.org/10.1186/s13229-021-00482-2.

Soriano, J.R., Daniels, N., Prinsen, J., Alaerts, K., 2020. Intranasal oxytocin enhances approach-related EEG frontal alpha asymmetry during engagement of direct eye contact. Brain Commun. 2, fcaa093. https://doi.org/10.1093/braincomms/fcaa093.

Sterne, J.A.C., Savović, J., Page, M.J., Elbers, R.G., Blencowe, N.S., Boutron, I., Cates, C. J., Cheng, H.Y., Corbett, M.S., Eldridge, S.M., Emberson, J.R., Hernán, M.A., Hopewell, S., Hróbjartsson, A., Junqueira, D.R., Jüni, P., Kirkham, J.J., Lasserson, T., Li, T., McAleenan, A., Reeves, B.C., Shepperd, S., Shrier, I., Stewart, L.A., Tilling, K., White, I.R., Whiting, P.F., Higgins, J.P.T., 2019. RoB 2: a revised tool for assessing risk of bias in randomised trials. Bmj 366 l4898. https://doi.org/10.1136/bmj.l4898.

Tager-Flusberg, H., Plesa Skwerer, D., Joseph, R.M., Brukilacchio, B., Decker, J., Eggleston, B., Meyer, S., Yoder, A., 2017. Conducting research with minimally verbal participants with autism spectrum disorder. Autism 21, 852–861. https://doi.org/10.1177/1362361316654605.

Viechtbauer, W., 2010. Conducting meta-analyses in r with the metafor package. J. Stat. Softw. 36, 1–48. https://doi.org/10.18637/jss.v036.i03.

Vivanti, G., Bent, C., Capes, K., Upson, S., Hudry, K., Dissanayake, C., Victorian, A.T., 2022. Characteristics of children on the autism spectrum who benefit the most from receiving intervention in inclusive versus specialised early childhood education settings. Autism Res. 15, 2200–2209. https://doi.org/10.1002/aur.2815.

Vivanti, G., Dissanayake, C., Zierhut, C., Rogers, S.J., Victorian, A.T., 2013. Brief report: predictors of outcomes in the early start Denver model delivered in a group setting. J. Autism Dev. Disord. 43, 1717–1724. https://doi.org/10.1007/s10803-012-1705-7.

Walsh, P., Elsabbagh, M., Bolton, P., Singh, I., 2011. In search of biomarkers for autism: scientific, social and ethical challenges. Nat. Rev. Neurosci. 12, 603–612. https://doi.org/10.1038/nrn3113.

Wang, Q., Wall, C.A., Barney, E.C., Bradshaw, J.L., Macari, S.L., Chawarska, K., Shic, F., 2020. Promoting social attention in 3-year-olds with ASD through gaze-contingent eye tracking. Autism Res. Off. J. Int. Soc. Autism Res. 13, 61–73. https://doi.org/10.1002/aur.2199.

Wass, S.V., Forssman, L., Leppänen, J., 2014. Robustness and precision: how data quality May influence key dependent variables in infant eye-tracker analyses. Infancy 19, 427–460.

Wei, Q., Dong, W., Yu, D., Wang, K., Yang, T., Xiao, Y., Long, D., Xiong, H., Chen, J., Xu, X., Li, T., 2024. Early identification of autism spectrum disorder based on machine learning with eye-tracking data. J. Affect Disord. 358, 326–334. https://doi.org/10.1016/j.jad.2024.04.049.

Wen, T.H., Cheng, A., Andreason, C., Zahiri, J., Xiao, Y., Xu, R., Bao, B., Courchesne, E., Barnes, C.C., Arias, S.J., Pierce, K., 2022. Large scale validation of an early-age eye-tracking biomarker of an autism spectrum disorder subtype. Sci. Rep. 12, 4253. https://doi.org/10.1038/s41598-022-08102-6.

Whittingham, K., McGlade, A., Kulasinghe, K., Mitchell, A.E., Heussler, H., Boyd, R.N., 2020. ENACT (ENvironmental enrichment for infants; parenting with Acceptance and Commitment Therapy): a randomised controlled trial of an innovative intervention for infants at risk of autism spectrum disorder. BMJ Open 10, e034315. https://doi.org/10.1136/bmjopen-2019-034315.

Williams, K., Wheeler, D.M., Silove, N., Hazell, P., 2010. Selective serotonin reuptake inhibitors (SSRIs) for autism spectrum disorders (ASD). Cochrane Database Syst. Rev. Cd004677. https://doi.org/10.1002/14651858.CD004677.pub2.

Yamasue, H., Okada, T., Munesue, T., Kuroda, M., Fujioka, T., Uno, Y., Matsumoto, K., Kuwabara, H., Mori, D., Okamoto, Y., 2020. Effect of intranasal oxytocin on the core social symptoms of autism spectrum disorder: a randomized clinical trial. Mol. Psychiatry 25, 1849–1858. https://doi.org/10.1038/s41380-018-0097-2.

Young, G.S., Merin, N., Rogers, S.J., Ozonoff, S., 2009. Gaze behavior and affect at 6 months: predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. Dev. Sci. 12, 798–814. https://doi.org/10.1111/j.1467-7687.2009.00833.x.

Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M.S., Saxena, S., Yusuf, A., Shih, A., Elsabbagh, M., 2022. Global prevalence of autism: a systematic review update. Autism Res. 15, 778–790. https://doi.org/10.1002/aur.2696.

Zhuang, H., Liang, Z., Ma, G., Qureshi, A., Ran, X., Feng, C., Liu, X., Yan, X., Shen, L., 2024. Autism spectrum disorder: pathogenesis, biomarker, and intervention therapy. MedComm 5, e497. https://doi.org/10.1002/mco2.497.