psychiatry international

MDPI

Systematic Review

Gender Bias in ASD Diagnostic and Screening Tools: A Systematic Review and Meta-Analysis

Ana Muiño Tato [1], Albert Marquès-Donoso [2,*] and Juan Carlos Sánchez-Huete [2]

[1] Department of Psychology, UNIE, C. de Arapiles, 14, 28015 Madrid, Spain; amuinot@professional.universidadunie.com
[2] Department of Social Sciences and Pedagogy, CES Don Bosco, C. María Auxiliadora, 9, 28040 Madrid, Spain; jcshuete@cesdonbosco.com
* Correspondence: amarques@cesdonbosco.com

Abstract

(1) Background: ASD diagnostic and screening instruments were historically developed on predominantly male samples, raising concerns about their differential validity across sexes. (2) Objective: This meta-analysis quantified the magnitude of gender bias in ASD diagnostic and screening tools and examined its consistency across instruments, age groups, and clinical contexts. (3) Method: A systematic search was conducted in APA PsycINFO, ERIC, Dialnet, and PsicoDoc (2015–2025) following PRISMA 2020 guidelines. Fourteen empirical studies were included (N = 55 to ~3,000,000). Effect sizes were computed as Fisher r-to-z transformed coefficients and pooled using a random-effects model. Heterogeneity was assessed via Cochran's Q and I^2; methodological quality was assessed via the CRF-QS. (4) Results: The pooled effect size was $\hat{\mu} = 0.29$ (95% CI: 0.11–0.47, z = 3.11, $p = 0.001$), indicating systematic bias favoring detection of externalizing, prototypically male profiles. Substantial heterogeneity was observed ($I^2 = 99.93\%$, $Q(13) = 11{,}417.61$, $p < 0.001$, $\tau^2 = 0.11$), with a prediction interval of −0.40 to 0.97. Funnel plot asymmetry suggested possible publication bias ($p = 0.01$). (5) Conclusions: ASD diagnostic instruments systematically underperform in identifying female presentations of autism. Bias operates across screening and formal evaluation stages, is not corrected by specialist assessment, and is associated with diagnostic delay and increased psychopathological burden. Revision of diagnostic tools and clinical training is urgently warranted.

Keywords: autism spectrum disorders; gender bias; diagnostic techniques; psychometric tests; sex differences in education; mental health

Academic Editor: Motohiro Okada

Received: 13 May 2026
Revised: 12 June 2026
Accepted: 25 June 2026
Published: 1 July 2026

1. Introduction

The diagnosis of autism spectrum disorder (ASD) relies fundamentally on standardized screening and clinical assessment instruments designed to operationalize the diagnostic criteria established in international classification manuals. These tools, which include parent-report questionnaires, structured clinical interviews, and standardized observational measures, have contributed to improved ASD detection over recent decades. However, their historical development has drawn predominantly on male samples and prototypical clinical profiles, raising relevant questions about their differential validity across sex and gender [1].

The epidemiological literature has consistently documented a marked disproportion in ASD diagnostic rates between males and females, with ratios ranging from 3:1 to 6:1

across different cultural and healthcare contexts [2]. Although these differences were long interpreted as reflecting a genuinely lower prevalence in females, accumulated evidence suggests that a substantial portion of this gap may be explained by systematic biases in identification and assessment processes, rather than by true differences in disorder prevalence [3,4].

It is important to acknowledge that the interpretation of sex-related differences in ASD diagnosis remains a matter of ongoing debate. Observed disparities in diagnostic rates, age at diagnosis, or instrument performance do not necessarily constitute evidence of diagnostic bias per se. Alternative explanations include genuine sex-based differences in prevalence, symptom expression, developmental trajectories, or biological risk factors. From this perspective, diagnostic inequities can only be interpreted as evidence of bias when differences arise from limitations in the assessment process rather than from true variations in autistic presentation. The present review adopts the position that repeated findings of delayed identification, differential instrument sensitivity, and the requirement for greater symptom burden among females may indicate systematic diagnostic disadvantage. Nevertheless, the distinction between sex differences and diagnostic bias should be considered when interpreting the findings.

A growing body of research has indicated that autistic women and girls tend to present with less externalizing phenotypic profiles, greater apparent social competence in structured contexts, and a higher reliance on social camouflaging strategies [5]. These characteristics are not incidental; they interact directly with the functioning of existing diagnostic tools, reducing their sensitivity to detect ASD when symptom presentation does not conform to the classic male phenotype upon which the instruments were built [6,7]. As a result, women typically require a higher symptom burden, more pronounced developmental delays, or additional comorbidities to meet diagnostic thresholds [8].

Despite growing theoretical recognition of gender bias in ASD diagnosis, the available empirical evidence remains fragmented and heterogeneous [9]. Existing studies vary widely in design, sample size, instruments employed, and reported metrics, making it difficult to draw integrated quantitative conclusions about the actual magnitude of bias associated with diagnostic tools. Advancing beyond narrative and systematic reviews to provide meta-analytic estimates capable of quantifying the effect of gender on the diagnostic performance of clinical and population-level instruments is therefore an important step forward.

The present meta-analysis focuses on examining sex- and gender-related differences in ASD screening and diagnostic instruments and evaluating the extent to which these differences may be consistent with the presence of diagnostic bias. Unlike prior work that addresses diagnostic bias in broad terms, this study adopts a focused approach centered on diagnostic tools as a key mechanism of inequity, enabling estimation of an aggregate effect size and evaluation of bias consistency across different contexts, age groups, and instruments.

A late or absent diagnosis limits early access to support, increases the risk of psychopathological comorbidity, and contributes to life trajectories marked by clinical invisibility, particularly in women and individuals assigned female at birth. A rigorous quantitative assessment of gender bias in diagnostic instruments therefore constitutes an essential step toward informing instrument revision, professional training, and the development of more equitable diagnostic models that are sensitive to the phenotypic diversity of ASD.

2. Materials and Methods

2.1. Study Design

A quantitative meta-analysis with a random-effects model was conducted to synthesize available empirical evidence on gender bias in ASD screening and diagnostic

instruments. The study was designed and reported in accordance with the recommendations of the PRISMA 2020 Statement for systematic reviews and meta-analyses [10]; the complete checklist is included as Supplementary Material. This systematic review and meta-analysis was retrospectively registered in PROSPERO: CRD420261424446.

The scope of this meta-analysis was specifically limited to the differential performance of diagnostic and screening instruments (e.g., parent-report questionnaires, clinical interviews, and standardized assessments) as a function of sex, excluding studies focused exclusively on prevalence or on clinical differences not linked to diagnostic tools. Using the PECO framework [11], the following research question guided this systematic review: Do current diagnostic and screening instruments adequately capture ASD in women and girls?

2.2. Literature Search Strategy

A systematic search was conducted in the APA PsycINFO, Dialnet, PsicoDoc, and ERIC databases, covering the period from 2015 to December 2025. The selection of these sources reflected a deliberate methodological decision: PsycINFO and ERIC provide broad coverage of high-impact international psychological and educational literature on ASD, while Dialnet and PsicoDoc allow for the inclusion of Spanish-language scientific output, whose diagnostic, cultural, and healthcare context presents specificities that the Anglophone literature does not always adequately capture. This combination responds to the goal of offering a synthesis that transcends the Anglocentric publication bias frequently noted in systematic reviews on ASD [12]. The search combined terms related to autism, gender, and diagnostic assessment using Boolean operators. The full search strategy is detailed in Table 1.

Table 1. Search Strategy (Boolean AND/OR).

Population (P)	Exposure (E)	Comparison (C)	Outcomes (O)
autism OR autistic OR "autism spectrum disorder" OR ASD OR neurodivergent	diagnos* OR "age at diagnosis" OR "diagnostic delay" OR "diagnostic pathway" OR identification OR detection OR screening OR assessment OR "camouflaging" OR masking OR compensation OR assimilation OR "sex differences" OR sex OR gender OR female OR women OR girls OR AFAB OR AMAB OR nonbinary OR transgender OR "gender diverse"	male OR female OR boys OR girls OR AFAB OR AMAB	prevalence OR epidemiology OR "age of diagnosis" OR "time to diagnosis" OR "waiting time" OR sensitivity OR specificity OR accuracy OR PPV OR NPV OR psychometric* OR "factor structure" OR invariance OR "measurement invariance" OR validity OR reliability OR "lived experience" OR qualitative OR interview* OR "thematic analysis" OR phenomenolog*

2.3. Eligibility Criteria

Studies were included if they met the following criteria: (1) quantitative empirical designs (observational, clinical, or population-based); (2) samples including children, adolescents, or adults assessed for ASD; (3) explicit use of ASD screening or diagnostic instruments (e.g., SCQ, ADOS, ADI-R, parent-report questionnaires, or structured clinical assessments); (4) analyses of sex-based differences in diagnosis probability, instrument performance, diagnostic thresholds, or diagnostic prediction; and (5) availability of sufficient statistics for effect size estimation or transformation.

Studies were excluded if they: (1) were qualitative designs, single-case studies, or narrative reviews; (2) focused exclusively on prevalence without analyzing the diagnostic

process; (3) did not disaggregate data by sex; (4) did not provide sufficient data for effect size calculation; or (5) were preprints not subjected to peer review.

Study Selection Process

The selection process was structured following a PECO framework adapted to the analysis of diagnostic bias:

- P (Population): Individuals assessed for suspected ASD or included in population-based studies using autism trait screening instruments.
- E (Exposure): Sex assigned at birth and/or gender, treated as a moderator of diagnostic performance; studies explicitly analyzing differences associated with gender socialization or camouflaging were included.
- C (Comparison): Contrasted group comparisons (girls vs. boys; AFAB vs. AMAB).
- (Outcomes): Indicators of diagnostic functioning, including diagnosis probability, instrument performance, diagnostic thresholds, or differential diagnostic prediction.

In the initial stage, the database search identified 142 articles. Following the removal of 11 duplicates, titles and abstracts of 131 articles were screened; 16 were excluded for failing to meet basic relevance criteria, yielding 115 studies for further analysis. Of these, 49 were excluded as systematic reviews or meta-analyses, leaving 66 articles for full-text eligibility review. At this stage, 52 studies were excluded for the following reasons: (1) 35 did not specifically investigate gender bias in diagnostic instruments; (2) 10 did not provide relevant data on screening outcomes or instrument utility; and (3) 7 did not meet the temporal criterion. A final sample of 14 articles was retained. The selection process is documented in the PRISMA flow diagram (see Figure 1). Title and abstract screening and full-text eligibility assessment were conducted by a single reviewer.

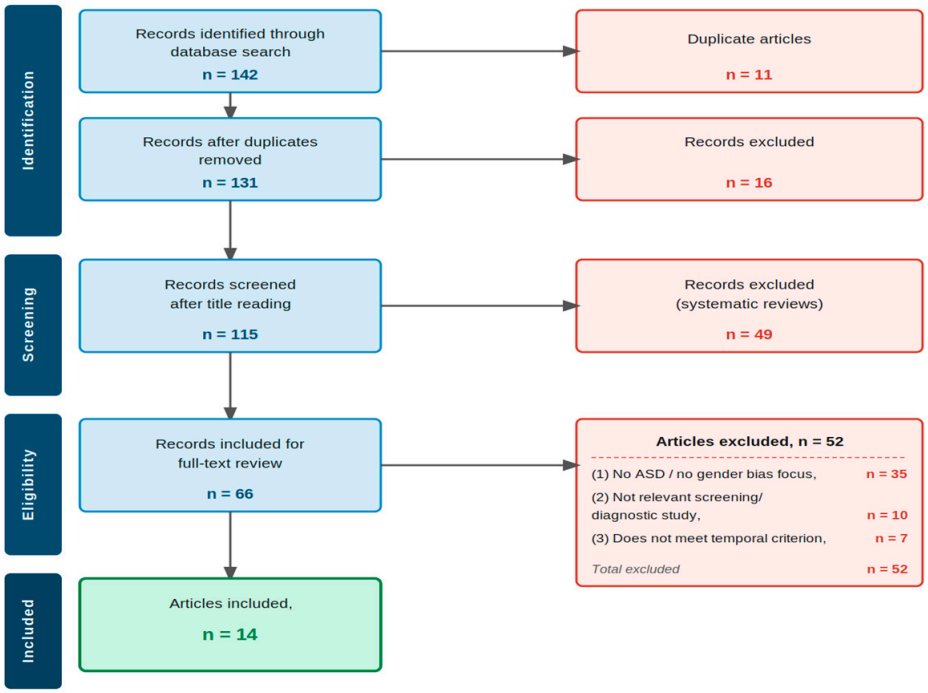

Figure 1. PRISMA Flow Diagram.

2.4. Data Extraction

The following data were systematically extracted from each included study: (1) author and year of publication; (2) country and study context; (3) study design; (4) total and sex-disaggregated sample size; (5) mean age; (6) diagnostic or screening instruments used;

(7) type of statistical analysis reported; and (8) relevant statistics (odds ratios, regression coefficients, proportion differences, confidence intervals). Data extraction was performed by a single reviewer using a standardized extraction form.

2.5. Methodological Quality Assessment

The methodological quality of included studies was assessed using the Critical Review Form—Quantitative Studies (CRF-QS), developed by Law et al. and widely used in systematic reviews in health sciences and psychology [13].

The CRF-QS provides a structured evaluation of key aspects of study design and methodological execution, including: clarity of study objectives, adequacy of design relative to the stated objective, description of the sample and selection criteria, validity and reliability of instruments used, appropriateness of statistical analysis, and interpretation of results (see Table 2). Methodological quality (CRF-QS) and risk of bias were assessed by a single reviewer.

Table 2. Methodological Quality of Included Studies (CRF-QS).

Study	Design	CRF-QS	Quality	Bias Risk	Rationale
Brian et al. [14]	Prospective longitudinal	17	Very good	Low	Blinded diagnosis, longitudinal follow-up, validated measures
Cage & Troxell-Whitman [15]	Cross-sectional observational (online survey) with mediation analysis	12	Moderate-high	Moderate	Disclosure measured via a single item, limiting construct depth. Methodological transparency, rigorous analysis, and conceptual relevance justify inclusion as a key theoretical support study for understanding mechanisms of diagnostic bias.
Duvekot et al. [16]	Multicenter observational	16	Very good	Low	Clear sex-based comparisons, standardized instruments
Hegemann et al. [17]	Population-based psychometric	18	Excellent	Low	Large N, EFA/CFA, factorial invariance tests
Kniola et al. [7]	Cross-sectional observational	16	Very good	Low	Large SPARK cohort, clear predictive analyses
Levante et al. [8]	Cross-sectional	14	Good	Uncertain	Parent self-reports, no diagnostic confirmation
McKinney et al. [9]	Cross-sectional observational	15	Good	Uncertain	Self-reported camouflaging, moderate sample
Morris & Campbell [18]	Population-based	16	Very good	Low	National survey, analyses by SES and gender
Parish-Morris [19]	Cross-sectional	15	Good	Uncertain	Reliance on traditional ADOS
Roman-Urrestarazu et al. [20]	Population-based	18	Excellent	Low	National registries, robust methodology
Rutherford et al. [21]	Quasi-experimental pre–post	16	Very good	Low	Objective indicators, structural improvement
Rutherford et al. [22]	Retrospective	15	Good	Uncertain	Real service data, referral bias
Smith et al. [23]	Observational with mediation	16	Very good	Low	Well-specified models, partial replication
Zahorodny et al. [24]	Population-based surveillance	16	Very good	Low	Multiple registries, active surveillance

2.6. Risk of Bias Assessment

The risk of bias of included studies was examined across dimensions relevant to observational and clinical studies, including: selection bias (sample source, clinical referral), measurement bias (use of instruments with potential sex-based sensitivity differences), confounding (control for variables such as age, cognitive level, and comorbidity), and reporting bias (selectivity in reported outcomes). This evaluation was integrated narratively into the interpretation of findings, with consideration of how specific methodological decisions may contribute to gender bias in diagnostic processes.

2.7. Statistical Analysis

All statistical analyses were conducted using Jamovi (Version 2.5.6) to ensure transparency and reproducibility. Effect sizes were calculated from correlation coefficients (r) transformed using Fisher's formula ($r \rightarrow z$), with the aim of quantifying the relationship between gender bias in females and ASD diagnostic and screening instruments. A random-effects model was employed, selected to account for the expected variability across studies and to provide a more precise estimate of the overall effect. Heterogeneity was considered across studies given methodological, contextual, and sampling differences, and was assessed using Cochran's Q statistic and the I^2 index as an estimate of the proportion of variance attributable to true heterogeneity. Results are presented using forest plots displaying individual effect sizes and the pooled effect with 95% confidence intervals. The dimensions of the analysis are summarized in Table 3.

Table 3. Meta-Analysis Dimensions.

Dimension	Description
Database search	APA PsycINFO, Dialnet, PsicoDoc, and ERIC
Date restrictions	2015–2025
Tau^2 estimate	$\tau^2 = 0.11$, indicating heterogeneity across studies
Cochran's Q test	Q ($p < 0.0001$), heterogeneity confirmed
I^2 statistic	$I^2 = 99.94\%$; variance explained by between-study heterogeneity
Outliers	No outlying studies identified
Prediction interval	Range: −0.40 to 0.97
Funnel plot asymmetry	Regression test ($p = 0.01$): possible publication bias; rank correlation non-significant

This study complies with all items of the PRISMA 2020 Checklist. The PRISMA flow diagram describes the process of study identification, selection, and inclusion, and the complete checklist is attached as Supplementary Material to ensure transparency and reproducibility.

Given the limited number of included studies (k = 14), a formal meta-regression with multiple moderators was not undertaken, as this would risk overfitting and unstable estimates (approximately ten studies per covariate are recommended). Exploratory subgroup analysis by study context was conducted instead (Section 3.3).

2.8. Use of Artificial Intelligence

Generative AI assistance (Claude Sonnet 4.5, Anthropic, San Francisco, CA, USA) was used during manuscript preparation solely to support writing tasks, including improving textual clarity, expositional coherence, and linguistic consistency across sections. All

research decisions, statistical analyses, empirical interpretations, and scientific content were produced exclusively by the authors.

3. Results

3.1. Included Studies

Fourteen studies were included in this meta-analysis, providing point estimates of effect size (ES) for gender bias in ASD diagnostic and screening instruments. Descriptive details, including year of publication, age range, symptom characteristics, and sample sizes, are summarized in Table 4.

Table 4. Descriptive Characteristics of Included Studies.

Author (Year)	Country/Context	Design	Sample (Total/by Sex)	Mean Age/Range	Diagnostic/ Screening Instruments	Relevant Statistics
Brian et al. [14]	Canada/High-risk family clinical cohort	Blinded prospective longitudinal	N = 67 (sex not reported by subgroup)	Follow-up to M ≈ 9.5 years	Blinded clinical assessment; standardized measures of autistic symptomatology, language, and cognition	Diagnostic stability analyses; kappa coefficient. Diagnostic stability 89.6%; κ = 0.76; ≈9% late diagnoses
Cage & Troxell-Whitman [15]	UK/Autistic adult community	Cross-sectional observational (online survey)	N = 180 (52% women; 42% men; 5% gender-diverse)	Adults	Camouflaging measures; autistic identity; disclosure	Significant mediation: identity → disclosure → ↓ camouflaging; competitive mediation
Duvekot et al. [16]	Netherlands/ Specialized clinical referral	Multicenter observational with logistic regression	N = 231 (ASD: 106 boys, 24 girls; non-ASD: 61 boys, 40 girls)	2.5–10 years	DDDI (short form); ADOS; CBCL; WISC/WPPSI/Bayley	Logistic regression with sex interaction terms. RRB × sex OR = 0.41; emotional problems × sex OR = 2.44
Hegemann et al. [17]	Norway/ Population cohort (MoBa)	Population-based psychometric study (EFA, CFA, factorial invariance)	EFA: 21,775; CFA: 21,674; Autism n = 636	8 years	SCQ (current version); national registries	Exploratory and confirmatory factor analysis; invariance tests. General sex invariance; qualitative differences in factor structure
Kniola et al. [7]	USA/SPARK national cohort	Cross-sectional observational with predictive models	N = 5946 autistic girls and women	Childhood–adulthood (verbal language)	SCQ; CBCL; parental developmental data	Multivariate logistic regression. Higher probability of positive screening with motor delays; implicit detection bias
Levante et al. [8]	Italy/Early general population	Cross-sectional with parent questionnaires	N = 361 (sex breakdown not detailed by subgroup)	18–36 months	Q-CHAT; ITSEA	Correlations and sex-based comparative analyses. Sex differences in autistic traits, externalizing behavior, and competence
McKinney et al. [9]	UK/Participatory school-based approach	Transdiagnostic cross-sectional observational	N = 119 girls (70 neurodivergent; 49 neurotypical)	M ≈ 12 years	CAT-Q-A; RCADS; ASC-ASD; SRS-2; Conners-3	Hierarchical regression; comparative analyses. Camouflaging predicts anxiety/depression; significant β independent of neurodevelopmental profile

Table 4. *Cont.*

Author (Year)	Country/Context	Design	Sample (Total/by Sex)	Mean Age/Range	Diagnostic/ Screening Instruments	Relevant Statistics
Morris & Campbell [18]	USA/National population survey	Survey-based observational	N = 224,401 (4.3% males; 1.2% females diagnosed)	3–17 years	National Survey of Children's Health	Proportion comparisons; logistic regression. Sex differences in diagnostic rate and age ($\Delta \approx 0.7$ years)
Parish-Morris [19]	USA (Children's Hospital of Philadelphia/University of Pennsylvania)	Experimental comparative study with eye-tracking (mixed design: between-groups ASD vs. TD; within-subjects by condition: rich vs. sparse social context)	N = 55 (ASD = 28; TD = 27). Sex: ASD (25 boys, 3 girls); TD (23 boys, 4 girls)	ASD: M = 27.04 years (21–49); TD: M = 28.19 years (20–48)	DSM-5 clinical diagnosis + ADOS-2; WASI-II (IQ); BAP-Q	Group × Condition interaction on face gaze: $\beta = -6.63$, $p = 0.01$. Main effect group: ASD < TD in face gaze ($\beta = -10.09$, $p = 0.03$). Main effect condition: greater face gaze in communicative context ($\beta = 12.81$, $p < 0.001$). Hand/object interaction: ASD gazed more at hands in social context ($\beta = 11.47$, $p = 0.01$). Global attention greater in rich social condition ($p = 0.0001$)
Roman-Urrestarazu et al. [20]	Chile/National education-health system	Observational with Bayesian analysis	≈3,000,000 children aged 6–18 years	School age	Health registries	Pre–post comparison: boy-to-girl ratio shifted from 5.6:1 to 2.7:1
Rutherford et al. [21]	UK/Clinical services	Quasi-experimental pre–post	Consecutive clinical samples	Childhood	Clinical record review	Sex-based comparisons across pre- and post-reform cohorts
Rutherford et al. [22]	UK/Clinical records	Retrospective observational	N = 150	Childhood–adulthood	Clinical record review	Significantly later diagnosis in girls
Smith et al. [23]	USA/Clinical and research setting	Observational with mediation	Clinical n = 1035 (22.9% AFAB); Research n = 128	Adolescence–adulthood	Age at diagnosis; anxiety/depression scales	Regression-based mediation models. AFAB → later diagnosis → ↑ anxiety/depression
Zahorodny et al. [24]	USA/Population surveillance	Longitudinal epidemiological	≈59,500 children	Childhood	Educational and health registries	Prevalence comparisons. Boy-to-girl ratio ≈ 5:1

The 14 studies included in the meta-analysis spanned a wide range of methodological and population contexts, including specialized clinical cohorts, large-scale population samples, and psychometric studies, with sample sizes ranging from intensive small-scale longitudinal designs (e.g., Brian et al. [14]) to population-level analyses based on millions of administrative records [20]. This diversity enabled examination of gender bias in the ASD diagnostic process across different developmental stages and levels of the diagnostic system, from early screening to formal clinical diagnosis.

Cross-sectional and longitudinal observational studies predominated in terms of methodological design, many incorporating advanced statistical analyses such as logistic regression models with sex interaction terms [7,16], mediation analyses [15,23] and psychometric studies with factorial invariance testing [17]. Several studies explicitly incorporated sex- or gender-based differential comparisons, analyzing either diagnosis probability, age at identification, or the predictive weight of specific symptom domains [18,21,22]. The most

frequently employed screening and diagnostic instruments were the Social Communication Questionnaire (SCQ), the Autism Diagnostic Observation Schedule (ADOS), parent-report behavioral and emotional regulation measures (CBCL, ITSEA, Q-CHAT), and clinical or population registries [8,19,24]. Some studies directly assessed the psychometric properties of these tools [17], while others examined their performance indirectly through diagnostic outcomes and clinical trajectories [7,16].

Across studies, systematic sex-based differences in the diagnostic process were consistently observed. In specialized clinical contexts, repetitive and restricted behaviors carried lower predictive weight for diagnosis in girls, whereas emotional and behavioral problems differentially increased diagnosis probability in this group [16]. In population-based screening studies, autistic girls and women required more prominent symptom manifestations or more marked developmental delays to exceed instrument thresholds, suggesting the existence of implicitly biased detection cutoffs [7,19].

Longitudinal and epidemiological studies provided convergent evidence of later diagnoses in girls, even when no differences were observed in the duration of the assessment process once clinical evaluation had been initiated [18,22]. These findings indicate that bias is not confined to the initial referral stage but also emerges within the formal diagnostic process itself. Mediation analyses further showed that diagnostic delay is indirectly associated with higher levels of anxiety and depression in individuals assigned female at birth, underscoring the clinical relevance of temporal bias [23]. The vast majority of effect sizes were positive and directionally consistent, indicating that ASD diagnostic instruments and processes systematically favor the identification of externalizing, prototypically male profiles [16,17,24].

3.2. Heterogeneity and Publication Bias

Analyses were conducted using Fisher's r-to-z transformed correlation coefficients as the outcome measure. A random-effects model was fitted to the data. The amount of heterogeneity (τ^2) was estimated alongside Cochran's Q test [25] and the I^2 statistic. Funnel plot asymmetry was assessed through both the rank correlation test and the regression test, using the standard error of observed outcomes as a predictor.

A total of k = 14 studies were included. Fisher r-to-z transformed correlation coefficients ranged from −0.22 to 0.99, with the majority of estimates positive (93%). The mean estimated Fisher-transformed correlation coefficient under the random-effects model was $\hat{\mu} = 0.29$ (95% CI: 0.11 to 0.47), differing significantly from zero (z = 3.11, $p = 0.001$) (see Figure 2). According to Cochran's Q test, the true outcomes appear to be heterogeneous (Q(13) = 11,417.61, $p < 0.0001$, $\tau^2 = 0.11$, $I^2 = 99.93\%$). The 95% prediction interval for the true outcomes ranged from −0.40 to 0.97; although the estimated mean outcome is positive, the true outcome may be negative in some individual studies. Examination of standardized residuals revealed that no study exceeded ±2.91 and therefore no outliers were identified in the context of this model. According to Cook's distances, no study was deemed unduly influential. Both the rank correlation test and the regression test indicated possible funnel plot asymmetry ($p = 0.002$ and $p = 0.01$, respectively). Taken together, the findings indicate that ASD diagnostic instruments and processes exhibit a systematic bias favoring the detection of externalizing, prototypically male profiles, penalizing more internalized or subtle presentations that are more common in girls ($\hat{\mu} = 0.29$).

Consistent with these findings, the aggregate effect sizes indicate that screening and diagnostic instruments do not operate equivalently for girls and boys, even when both groups have been referred for clinical evaluation or present with comparable diagnostic suspicion. This finding reinforces the view that diagnostic inequities cannot be explained

solely by differences in prevalence or delays in referral, but are actively produced within the assessment process itself.

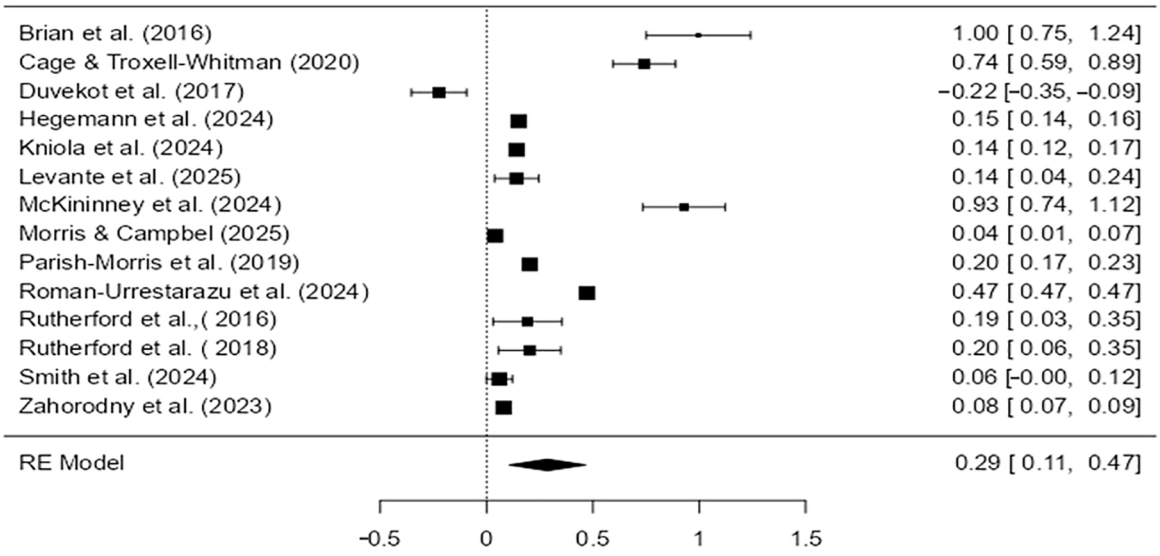

Figure 2. Forest Plot [7–9,14–24]. Squares represent individual study effect sizes, with size proportional to study weight in the meta-analysis; the diamond represents the pooled random-effects estimate.

3.3. Exploratory Subgroup Analysis by Study Context

Given the substantial heterogeneity observed in the overall model, an exploratory subgroup analysis was conducted according to study context, distinguishing between clinical and population-based studies (see Table 5). Clinical studies yielded a positive but non-significant pooled effect size (Fisher's z = 0.237, SE = 0.197, 95% CI: −0.150 to 0.624, p = 0.231), with substantial residual heterogeneity (τ^2 = 0.188, I^2 = 94.98%, Q = 79.68). Population-based studies also showed a positive pooled effect size, which reached statistical significance (Fisher's z = 0.172, SE = 0.064, 95% CI: 0.046 to 0.297, p = 0.007), although heterogeneity remained extremely high (τ^2 = 0.024, I^2 = 99.95%, Q = 10,760.29).

Table 5. Pooled Effect Sizes and Heterogeneity by Study Context (Clinical vs. Population-Based).

Study Context	k	Fisher's z	95% CI	*p*	τ^2	I^2 (%)	Q
Clinical	5	0.237	−0.150 to 0.624	0.231	0.188	94.98	79.68
Population-based	6	0.172	0.046 to 0.297	0.007	0.024	99.95	10,760.29
Between-group difference	—	0.057	−0.301 to 0.414	0.757	—	—	—

The between-subgroup comparison did not indicate a statistically significant difference between clinical and population-based contexts (β = 0.057, SE = 0.182, 95% CI: −0.301 to 0.414, p = 0.757). These findings suggest that study context alone does not account for the heterogeneity observed in the overall model. Rather than supporting a uniform effect across settings, the results indicate that sex- and gender-related diagnostic disparities vary considerably across studies and are likely influenced by additional methodological, developmental, and instrument-specific factors.

4. Discussion

4.1. Clinical Interpretation

The findings suggest that the observed diagnostic disparities cannot be attributed solely to clinical application errors or to contextual factors occurring prior to evaluation, but rather reflect deeper issues related to the design and conceptualisation of diagnostic instruments themselves.

The results are consistent with the construct bias hypothesis, according to which the instruments function correctly in relation to the construct they measure, but that construct is defined in a partial and androcentric manner [3].

Widely used instruments such as the ADOS, ADI-R, and SCQ were historically developed and validated on predominantly male samples, in a context in which autism was conceptualized on the basis of prototypical profiles associated with male children exhibiting clearly observable behavioral manifestations [4]. This history translates into a disproportionate emphasis on externalizing, disruptive, or socially conspicuous behaviors, as well as on explicit forms of restricted interests and social difficulties that are more readily detectable in boys than in girls [26].

As a consequence, these instruments show reduced sensitivity to internalizing manifestations of autism, which are more prevalent in girls and women, such as emotional distress, social anxiety, interpersonal hypervigilance, and social overcompensation. Restricted interests expressed through socially normative content (e.g., literature, animals, interpersonal relationships, or cultural figures) tend to be undervalued or interpreted as compatible with typical development, thereby reducing their diagnostic weight despite meeting comparable criteria for intensity, rigidity, or centrality to those observed in boys [27].

This pattern has been empirically documented in recent observational studies. Duvekot et al. demonstrated that parent-reported repetitive and restricted behaviors carry lower predictive value for diagnosis in girls, whereas emotional and behavioral problems increase diagnosis probability in this group, suggesting that girls must present additional or atypical symptomatology to be clinically recognized [16]. These findings align with those in the broader literature, in which the differential weighting of specific domains contributes substantially to aggregate diagnostic bias [1].

Kniola et al. provide large-scale evidence that autistic girls and women require greater symptom severity or more pronounced developmental milestone delays to exceed the screening thresholds of instruments such as the SCQ [7]. This phenomenon points to the existence of implicitly elevated diagnostic thresholds that do not correspond to a lower presence of autism, but rather to an inadequate fit of the instrument for detecting less externalizing or more compensated profiles [28].

Findings from Rutherford et al. show that even in specialized clinical services, girls are referred and diagnosed later than boys, without any observed differences in the duration of the assessment process once initiated [22]. This finding is particularly relevant, as it suggests that bias does not occur solely at the point of initial access but persists within formal assessment itself, reinforcing the interpretation that instruments are not optimized to identify female ASD profiles. Similar patterns have been reported by Goldblum et al., who observed diagnostic delays in certain subgroups of girls and women, particularly among non-Hispanic groups including White, Black, and Asian individuals [2].

The exploratory subgroup analysis suggests that sex- and gender-related diagnostic disparities are not confined to a single study context. However, the persistence of substantial heterogeneity within both clinical and population-based subgroups indicates that these disparities should not be interpreted as operating uniformly across settings. The evidence is therefore better understood as suggesting context-dependent diagnostic inequities rather than a single overarching effect of equivalent magnitude across all diagnostic systems.

4.2. The Role of Camouflaging

Female socialization promotes, from early developmental stages, behaviors oriented toward interpersonal emotional regulation and adaptation to implicit social norms. In autistic girls and women, this process can result in reduced behavioral visibility of core ASD symptoms, particularly in structured or high social-demand contexts such as clinical assessment. The likelihood of false negatives consequently increases, particularly when instruments prioritize the direct observation of explicit social deficits or repetitive behaviors. The literature has conceptualized this phenomenon under the term camouflaging or masking, defined as the conscious or unconscious use of strategies to conceal, compensate for, or normalize autistic traits in social environments [5,29,30].

The findings of the present study support the need to move beyond a conception of camouflaging as a purely individual variable and to understand it as an interactive factor that modulates the performance of diagnostic instruments. Hull et al. demonstrated that autistic women report significantly higher levels of camouflaging than men, even after controlling for symptom severity, and that these strategies are associated with greater psychological exhaustion and poorer mental health [30]. From a diagnostic perspective, these findings imply that instruments based on direct behavioral observation may underestimate symptomatology when the individual employs effective compensatory strategies, particularly in brief, structured assessment sessions.

Tien et al. have proposed theoretical models in which camouflaging acts as a moderator between the phenotypic expression of autism and its clinical detection, particularly in women with average or high cognitive profiles [1]. These models are broadly consistent with the findings of the present meta-analysis, in which sex- and gender-related diagnostic disparities were observed across both clinical and population-based contexts. However, given the substantial heterogeneity observed, the present findings should not be interpreted as demonstrating a uniform mechanism across all assessment settings. Friedman et al. have similarly shown that many autistic adult women describe diagnostic trajectories marked by years of forced social adaptation, reframing of their difficulties as anxiety or emotional problems, and an internalization of distress that impedes clinical recognition of ASD. This pattern is consistent with the meta-analytic findings, in which internalizing domains carry greater diagnostic weight in girls only when they reach clinically significant levels, while more subtle manifestations go undetected [5].

This pattern aligns with findings from population-based epidemiological studies such as those of Zahorodny et al., which document significant discrepancies between estimated ASD prevalence and the proportion of individuals with a confirmed diagnosis, particularly among women and those with average or high cognitive profiles [25]. These studies suggest that a substantial proportion of adolescents and adults retrospectively identified as autistic had not received a diagnosis in childhood, despite having had contact with educational or healthcare systems. The available evidence suggests that diagnostic disparities may persist at the stage of clinical assessment, even when gold-standard instruments such as the ADOS or ADI-R are employed. However, the number of studies directly addressing this issue remains limited and the present findings should therefore be interpreted with caution. Although these protocols offer a more in-depth and multimodal evaluation, they remain anchored in observational criteria and prototypical behavioral exemplars that may not adequately capture the autistic phenomenology of girls and women, particularly when effective social compensatory strategies are present.

The available evidence suggests that diagnostic disparities may persist even within specialist assessment settings. However, the number of studies directly addressing this issue remains limited, and the present findings should therefore be interpreted cautiously. Further research is needed to determine the extent to which specialist evaluation mitigates

or maintains sex- and gender-related differences in ASD identification. Screening limits who reaches the clinic, and the clinic in turn preferentially validates profiles that conform to a historically male phenotypic model of autism. As a result, girls who do access evaluation typically do so after a greater degree of functional impairment, emotional comorbidity, or adaptive failure, which introduces an additional bias into clinical samples and reinforces the false impression of lower female prevalence.

4.3. Limitations

A primary limitation of this meta-analysis is that its focus on sex-based differences in the performance of diagnostic instruments precludes full capture of the qualitative and dynamic processes involved in clinical encounters, such as the evaluator–patient interaction, professional expectations, and implicit decision-making during information integration. These factors, which are extensively documented in the qualitative literature, likely contribute to diagnostic bias but fall outside the scope of the quantitative designs included here.

Although a systematic assessment of methodological quality and risk of bias was conducted (CRF-QS), the total number of studies meeting strict eligibility criteria remains limited, particularly for specific age subgroups or individual instruments. This reflects a structural gap in the existing research base and reinforces the need for studies explicitly designed to evaluate gender bias from the early stages of diagnostic instrument development. Study selection, data extraction, and quality assessment were carried out by a single reviewer rather than independently by two or more reviewers. This may increase the risk of selection and extraction errors and is acknowledged as a methodological limitation of the review process.

The literature search was restricted to four databases, which may have limited retrieval of studies indexed exclusively in MEDLINE or Web of Science. However, this decision reflects a deliberate approach oriented toward capturing both high-impact international output and Spanish-language scientific literature, which is frequently absent from meta-analyses on ASD with an exclusively Anglophone scope.

An important conceptual limitation concerns the distinction between observed sex differences and diagnostic bias. The studies included in this meta-analysis primarily assessed differences in diagnostic outcomes, screening performance, age at identification, and symptom thresholds across males and females. Although these patterns are consistent with the hypothesis of diagnostic bias, they do not independently demonstrate that all observed differences result from inequitable assessment procedures. Some variability may reflect genuine differences in symptom expression, developmental pathways, or prevalence. Consequently, the findings should be interpreted as evidence of systematic diagnostic disparities that are compatible with, but do not by themselves definitively prove, the existence of diagnostic bias.

A further limitation is that several included studies examined symptom severity, screening scores, or diagnostic thresholds rather than categorical diagnostic outcomes. Consequently, the pooled effect size reflects broader diagnostic disparities and instrument performance differences rather than diagnostic status alone. The distinction between severity-related effects and diagnostic outcomes should therefore be considered when interpreting the magnitude of the observed associations. Furthermore, not all included studies directly evaluated the performance of ASD screening or diagnostic instruments. For example, Parish-Morris et al. examined behavioural processes related to the visibility and recognition of autistic characteristics rather than the psychometric performance of a diagnostic tool. Although such studies may provide indirect evidence relevant to diagnostic

identification, their inclusion introduces additional conceptual heterogeneity that should be considered when interpreting the pooled estimates [18].

Although one included study reported the participation of gender-diverse individuals [15], the overwhelming majority of studies analysed only binary male–female comparisons. As a result, the present findings cannot be generalised to autistic individuals with diverse gender identities. Future research should explicitly examine diagnostic experiences among transgender, non-binary, and gender-diverse autistic populations.

Finally, the extreme heterogeneity observed (I^2 = 99.93%) limits the extent to which a single pooled estimate can be considered representative across all diagnostic contexts. An exploratory subgroup analysis by study context (clinical vs. population-based) was conducted, but substantial residual heterogeneity remained within both subgroups. A formal meta-regression incorporating the available moderators was not undertaken, as the number of included studies (k = 14) is well below the threshold generally recommended for stable estimation (approximately ten studies per covariate), and would have resulted in an overfitted model. Stratified subgroup analyses across the remaining moderators were likewise precluded by the distribution of the evidence: only a single study fell clearly within the adolescent age range; several studies did not employ a standardized ASD screening or diagnostic instrument as their primary measure, preventing a reliable screening-versus-diagnostic classification; and longitudinal designs were too few relative to cross-sectional ones to support stable subgroup estimates. These constraints indicate that the observed effect should be interpreted as evidence of context-dependent diagnostic disparities rather than a uniform bias of equivalent magnitude across settings, and they identify age, instrument type, and study design as priority moderators for future, adequately powered syntheses.

4.4. Methodological Implications and Future Research Directions

The findings of this meta-analysis underscore the urgency of advancing toward gender-sensitive diagnostic assessment models, both in the development of new instruments and in the critical revision of existing ones. Evidence of construct bias suggests that future tools should incorporate domains less dependent on direct behavioral observation and more attentive to internal processes, adaptive costs, and compensatory strategies.

A priority research direction involves differential item and threshold analyses by sex and gender, not with the aim of fragmenting the diagnosis, but of improving precision and reducing systematic errors. Studies exploring the factorial invariance of diagnostic instruments can provide critical information about which dimensions function differentially and at which developmental stages.

The results also point to the need for longitudinal designs capable of examining how diagnostic bias is constructed and transforms over time. Understanding at which developmental moments gender-based discrepancies are amplified can inform preventive strategies and early adjustments in the assessment process.

An additional relevant direction is the systematic integration of moderating variables such as cognitive level, emotional comorbidity, and sociocultural context. Far from being mere confounders, these variables form part of the real-world context in which diagnosis occurs and should be explicitly incorporated into analytical models. Meta-analytic approaches incorporating multilevel models or meta-regression could contribute to a more refined understanding of these interactions.

5. Conclusions

This study contributes to shifting focus from a narrative centered on the invisibility of autistic women toward a more precise understanding of the role those diagnostic systems

themselves play in the production of inequities. Addressing gender bias in ASD diagnosis does not require only earlier detection; it requires diagnosing differently, incorporating assessment models that recognize the heterogeneity of the spectrum and challenge the implicit assumptions that have historically guided its definition.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/psychiatryint7040145/s1. Table S1: Completed PRISMA 2020 Checklist.

Author Contributions: Conceptualization, A.M.T. and A.M.-D.; methodology, A.M.T. and A.M.-D.; formal analysis, A.M.T.; investigation, A.M.T.; data curation, A.M.T.; writing—original draft preparation, A.M.T.; writing—review and editing, A.M.T., A.M.-D. and J.C.S.-H.; translation and linguistic adaptation, A.M.-D.; methodological adequacy and final version preparation, A.M.-D.; visualization, A.M.T.; supervision, A.M.-D.; statistical review, A.M.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Centro de Enseñanza Superior Don Bosco. The APC was funded by Centro de Enseñanza Superior Don Bosco (Number 003-2026).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings of this study (extracted effect sizes and coding sheets) are available on request from the corresponding author. The original datasets analyzed derive from previously published studies; readers are referred to the original sources for primary data access.

Acknowledgments: During the preparation of this manuscript, the authors used Claude Sonnet 4.5 (Anthropic) for the purposes of reviewing and improving the clarity and expositional coherence of the text, and refining the linguistic formulation of certain sections. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest. The funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

ADI-R	Autism Diagnostic Interview—Revised
ADHD	Attention-Deficit/Hyperactivity Disorder
ADOS/ADOS-2	Autism Diagnostic Observation Schedule (2nd edition)
AFAB	Assigned Female At Birth
AMAB	Assigned Male At Birth
APA	American Psychological Association
ASD	Autism Spectrum Disorder
ASC-ASD	Anxiety Scale for Children—Autism Spectrum Disorder
BAP-Q	Broader Autism Phenotype Questionnaire
CAT-Q/CAT-Q-A	Camouflaging Autistic Traits Questionnaire (Adult version)
CBCL	Child Behavior Checklist
CFA	Confirmatory Factor Analysis
CI	Confidence Interval
CRF-QS	Critical Review Form—Quantitative Studies
DDDI	Developmental, Dimensional and Diagnostic Interview
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
EFA	Exploratory Factor Analysis

ERIC	Education Resources Information Center
ES	Effect Size
I^2	I-squared statistic (index of between-study heterogeneity)
IQ	Intelligence Quotient
ITSEA	Infant–Toddler Social and Emotional Assessment
NPV	Negative Predictive Value
OR	Odds Ratio
PECO	Population, Exposure, Comparison, Outcomes (framework)
PPV	Positive Predictive Value
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
Q-CHAT	Quantitative Checklist for Autism in Toddlers
RCADS	Revised Children's Anxiety and Depression Scale
RRB	Restricted and Repetitive Behaviors
SCQ	Social Communication Questionnaire
SES	Socioeconomic Status
SPARK	Simons Foundation Powering Autism Research for Knowledge
SRS-2	Social Responsiveness Scale, 2nd Edition
TD	Typically Developing
UNED	Universidad Nacional de Educación a Distancia
UNIE	Universidad Internacional de la Empresa
WASI-II	Wechsler Abbreviated Scale of Intelligence, 2nd Edition
WISC	Wechsler Intelligence Scale for Children
WPPSI	Wechsler Preschool and Primary Scale of Intelligence

References

1. Tien, I.; Pearson, A.; Sozeri, S.; Seers, K. "Only boys can be autistic": Gender stereotypes and socialization in the diagnostic journey. *Autism* **2025**. [CrossRef]
2. Goldblum, J.E.; McFayden, T.C.; Bristol, S.; Putnam, O.C.; Wylie, A.; Harrop, C. Autism prevalence and the intersectionality of assigned sex at birth, race, and ethnicity on age of diagnosis. *J. Autism Dev. Disord.* **2024**, *54*, 3777–3791. [CrossRef] [PubMed]
3. Probol, T.; Mieskes, M. A reproducible dataset of autistic speech from social media: Gender representation and linguistic features. *Proc. ACL* **2025**. [CrossRef]
4. Rea, C.J.; Lesch, K.; Hernandez, B.; Sprecher, E.; Hauptman, M.; Epee-Bounya, A.; Wilson, K.; Shah, S.N. Child and family characteristics associated with age of autism spectrum disorder diagnosis in a primary care population. *J. Dev. Behav. Pediatr.* **2025**, *46*, e450–e455. [CrossRef] [PubMed]
5. Friedman, A.; Paltoglou, A.; Sorte, R. A qualitative exploration of the experiences of self-diagnosed autistic women and gender-diverse individuals who are not pursuing an autism diagnosis. *Autism Adulthood* **2024**, *2*, 27546330241307828. [CrossRef]
6. Kentrou, V.; de Veld, D.M.J.; Mataw, K.J.K.; Begeer, S. Delayed autism spectrum disorder recognition in children and adolescents previously diagnosed with ADHD. *Autism* **2019**, *23*, 1065–1072. [CrossRef] [PubMed]
7. Kniola, A.; Ludwig, N.N.; Singh, V.; Bradley, C.; Carpenter, L.; Dillon, E.F.; Kanne, S.; Kim, S.H.; Parish-Morris, J.; Snyder, L.G.; et al. Delayed milestones and demographic factors relate to autism screening accuracy in verbally fluent autistic girls. *J. Autism Dev. Disord.* **2024**, *56*, 547–559. [CrossRef] [PubMed]
8. Levante, A.; Martis, C.; Duma, L.; Lecciso, F. Autism traits, emotional regulation, and competence in toddlers: Sex differences in early development. *Infant Behav. Dev.* **2025**, *117*, 202679. [CrossRef]
9. McKinney, A.; O'Brien, S.; Maybin, J.A.; Chan, S.W.Y.; Richer, S.; Rhodes, S. Camouflaging in neurodivergent and neurotypical girls at the transition to adolescence and its relationship to mental health: A participatory methods research study. *J. Child Adolesc. Couns.* **2024**, *4*, e12294. [CrossRef] [PubMed]
10. 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.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef] [PubMed]
11. Petticrew, M.; Roberts, H. *Systematic Reviews in the Social Sciences: A Practical Guide*; Blackwell Publishing: Oxford, UK, 2006. [CrossRef]
12. Elsabbagh, M.; Divan, G.; Koh, Y.J.; Kim, Y.S.; Kauchali, S.; Marcín, C.; Montiel-Nava, C.; Patel, V.; Paula, C.S.; Wang, C.; et al. Global prevalence of autism and other pervasive developmental disorders. *Autism Res.* **2012**, *5*, 160–179. [CrossRef] [PubMed]

13. Law, M.; Stewart, D.; Pollock, M.; Letts, L.; Bosch, J.; Westmorland, M. *Guidelines for Critical Review Form: Quantitative Studies*; McMaster University Occupational Therapy Evidence-Based Practice Research Group: Hamilton, ON, Canada, 1998.
14. Brian, J.; Bryson, S.E.; Smith, I.M.; Roberts, W.; Roncadin, C.; Szatmari, P.; Zwaigenbaum, L. Stability and change in autism spectrum disorder diagnosis from age 3 to middle childhood in a high-risk sibling cohort. *Autism* **2016**, *20*, 888–892. [CrossRef] [PubMed]
15. Cage, E.; Troxell-Whitman, Z. Understanding the relationships between autistic identity, disclosure, and camouflaging. *Autism Adulthood* **2020**, *2*, 334–338. [CrossRef] [PubMed]
16. Duvekot, J.; van der Ende, J.; Verhulst, F.C.; Slappendel, G.; van Daalen, E.; Maras, A.; Greaves-Lord, K. Factors influencing the probability of a diagnosis of autism spectrum disorder in girls versus boys. *Autism* **2017**, *21*, 646–658. [CrossRef] [PubMed]
17. Hegemann, L.; Askeland, R.B.; Valand, S.B.; Øyen, A.S.; Schjølberg, S.; Bal, V.H.; Bishop, S.L.; Stoltenberg, C.; von Soest, T.; Hannigan, L.J.; et al. Measuring autism-associated traits in the general population: Factor structure and measurement invariance across sex and diagnosis status of the Social Communication Questionnaire. *Autism* **2024**, *28*, 2105–2119. [CrossRef] [PubMed]
18. Morris, J.; Campbell, M. Investigating the Disparities in Autism Spectrum Disorder Diagnoses. Poster presented at the Research and Schol-arship Conference (RASC), Western Carolina University, Cullowhee, NC, USA. 2025. Available online: https://affiliate.wcu.edu/rasc24/wp-content/uploads/sites/404/2024/03/Trey.Morris.RASC_.Poster.pdf (accessed on 1 October 2025).
19. Parish-Morris, J.; Pallathra, A.A.; Ferguson, E.; Maddox, B.B.; Brodkin, E.S.; Schultz, R.T. Adaptation to different communicative contexts: An eye tracking study of autistic adults. *J. Neurodev. Disord.* **2019**, *11*, 5. [CrossRef] [PubMed]
20. Román-Urrestarazu, A.; van Kessel, R.; Matthews, F.; Allison, C.; Brayne, C.; Baron-Cohen, S. Autism prevalence in Chile: A national Bayesian analysis of health and education records. *Autism Res.* **2024**, *17*, 345–357. [CrossRef] [PubMed]
21. Rutherford, M.; Burns, M.; Gray, D.; Bremner, L.; Clegg, S.; Russell, L.; Smith, C.; O'Hare, A. Improving efficiency and quality of the children's ASD diagnostic pathway: Lessons learned from practice. *J. Autism Dev. Disord.* **2018**, *48*, 1579–1595. [CrossRef] [PubMed]
22. Rutherford, M.; McKenzie, K.; Johnson, T.; Catchpole, C.; O'Hare, A.; McClure, I.; Forsyth, K.; McCartney, D.; Murray, A. Gender ratio, age of diagnosis, and assessment duration in autism spectrum disorder. *Autism* **2016**, *20*, 628–634. [CrossRef] [PubMed]
23. Smith, J.V.; McQuaid, G.A.; Wallace, G.L.; Neuhaus, E.; Lopez, A.; Ratto, A.B.; Jack, A.; Khuu, A.; Webb, S.J.; Verbalis, A.; et al. Time is of the essence: Age at autism diagnosis, sex assigned at birth, and psychopathology. *Autism* **2024**, *28*, 2909–2922. [CrossRef] [PubMed]
24. Zahorodny, W.; Shenouda, J.; Sidwell, K.; Verile, M.G.; Cruz Alvarez, C.; Fusco, A.; Mars, A.; Waale, M.; Gleeson, T.; Burack, G.; et al. Prevalence and characteristics of adolescents with autism spectrum disorder in the New York–New Jersey metropolitan area. *J. Autism Dev. Disord.* **2023**, *55*, 2043–2049. [CrossRef] [PubMed]
25. Cochran, W.G. The Combination of Estimates from Different Experiments. *Biometrics* **1954**, *10*, 101–129. [CrossRef] [PubMed]
26. Salomon, C.; Heinz, K.; Aronson-Ramos, J.; Wall, D.P. An analysis of the real world performance of an artificial intelligence based autism diagnostic. *Sci. Rep.* **2025**, *15*, 29503. [CrossRef] [PubMed]
27. Surgent, O.; Andrews, D.S.; Lee, J.K.; Boyle, J.; Dakopolos, A.; Miller, M.; Ozonoff, S.; Rogers, S.J.; Solomon, M.; Amaral, D.G.; et al. Sex differences in the striatal contributions to longitudinal fine motor development in autistic children. *Biol. Psychiatry* **2025**, *97*, 1150–1162. [CrossRef] [PubMed]
28. Peterson, T.; Dodson, J.; Sherwin, R.; Strale, F., Jr. Comparative effects of applied behavior analysis on male and female individuals with autism spectrum disorder. *Cureus* **2024**, *16*, e59802. [CrossRef] [PubMed]
29. Čolić, M.; Dababnah, S.; Garbarino, N.; Betz, G. Parental experiences raising children with autism spectrum disorder in Eastern Europe: A scoping review. *Int. J. Dev. Disabil.* **2019**, *68*, 1–13. [CrossRef] [PubMed]
30. Hull, L.; Mandy, W.; Lai, M.C.; Baron-Cohen, S.; Allison, C.; Smith, P.; Petrides, K.V. Development and validation of the Camouflaging Autistic Traits Questionnaire (CAT-Q). *J. Autism Dev. Disord.* **2019**, *49*, 819–833. [PubMed]