Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s10803-025-07008-2

ORIGINAL ARTICLE

Psychometric Properties of Measuring Tools for Depression in Autistic Youths: A Systematic Review

Héloïse Young[1] · Hélène Lahaye[2,3] · Jean-Marc Guilé[2,3] · Cora Cravero[4,5] · Angèle Consoli[4,6] · David Cohen[4,7] · Real Labelle[8,9] · Xavier Benarous[4,6]

Accepted: 25 July 2025

Abstract

While autistic individuals are 4 times more likely to have depression compared to non-autistic individuals, depression remained largely undetected in this group resulting in significant unmet needs. A systematic literature review was conducted to determine the psychometric properties of measuring tools for depressive symptoms in autistic children and adolescents. The review followed the PRISMA 2020 guideline (PROSPERO: CRD42023423377) based on the search of several databases from 1980 until November 2024. The COnsensus-based Standards for the Selection of health status Measurement INstruments (COSMIN) checklist was used to assess for each tool: internal consistency, reliability, measurement error, content validity, structural validity, hypothesis testing, criterion validity, responsiveness to change and cross-cultural validity. The review found 15 empirical studies evaluating the properties of twelve measuring tools for depression in autistic youths. The validities of standards screening scales were low, but better for questionnaires with items focusing on behavioral aspects of depression, i.e., the Patient Health Questionnaire-9 (PHQ-9) and the Hospital Anxiety and Depression Scale (HADS). The evidence supporting the quality of the instrument was much better for those developed for subjects with neurodevelopmental divergences, such as The Evaluation of Depressive symptoms in Autism (EDA), with five domains of psychometric properties rated as strong or moderate. To determine the optimal approach for the use of consensual instruments, further research should include more individuals with co-occurring neurodevelopmental conditions, a risk factor for a more resistant form of depression, and examine the instruments' capacity to detect clinically significant changes during an intervention.

Keywords Mood disorders · Neurodevelopmental disorders · Irritability · Measuring instrument · Outcome

✉ Xavier Benarous
xavier.benarous@aphp.fr

1 Department of Child and Adolescent Psychopathology, Rouen University Hospital, Rouen, France

2 Department of Child and Adolescent Psychopathology, Amiens University Hospital, Amiens, France

3 INSERM Unit U1105 Research Group for Analysis of the Multimodal Cerebral Function, University of Picardy Jules Verne, Amiens, France

4 Department of Child and Adolescent Psychiatry, Pitié-Salpêtrière University Hospital, APHP.Sorbonne Université, 83 boulevard de l'Hôpital, 75013 Paris, France

5 Interdepartmental Mobile Unit for Complex Situations in Autism (UMI 75-92), Elan Retrouvé Foundation, Paris, France

6 Sorbonne Université, INSERM, UMR-S 1136 Institut Pierre Louis d'Epidémiologie et de Santé Publique (iPLesp), Equipe Epidémiologie Sociale, Santé Mentale, Addictions (ESSMA), 75012 Paris, France

7 CNRS UMR 7222, Institute for Intelligent Systems and Robotics, Sorbonne University, Paris, France

8 Department of Psychology, Université du Québec à Montréal, Montréal, Québec, Canada

9 Centre for Research and Intervention On Suicide, Ethical Issues and End-of-Life Practices, Université du Québec à Montréal, Montréal, Québec, Canada

Published online: 20 August 2025

Introduction

Autism represents a range of neurodevelopmental differences characterized by specificities in social interaction and communication, specialized interests, repetitive behaviors, and atypical sensory processing. In the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR), autism was operationalized as a spectrum with large variability in terms of characteristics, adaptive skills, and cognitive functioning (American Psychiatric Association, 2022). Co-occurring neurodivergence and mental health issues largely contribute to this heterogeneity, with almost three-quarters of autistic youths matching the criterion for at least one associated psychiatric disorder (Astle et al., 2022; Hossain et al., 2020; Lai et al., 2019).

The lifetime prevalence of depressive disorders (i.e., major depressive disorder, persistent depressive disorder) in autistic individuals at adulthood is estimated between 9 and 54% (Lai et al., 2019; Leyfer et al., 2006), with a pooled lifetime prevalence four times higher than in non-autistic individuals (Hudson et al., 2019). In a subgroup analysis of their meta-analysis, Hudson et al. (2019) reported a pooled lifetime prevalence of depression of 7.7%, 95% CI 4.7–12.4 in participants under 18. Studies conducted in pediatric samples showed that the increased risk of suffering from depression in autistic individuals at these ages results from a combination of individual factors and environmental stressors, in particular, peer victimization and stressful life experiences due to social barriers and stigma (Benarous et al., 2023; Farhat et al., 2023; Gussin et al., 2024; Pezzimenti et al., 2019; Rai et al., 2018; Thapar et al., 2023). The comparison of depression features between autistic and non-autistic youths supports higher severity of depressive symptoms, suicidal behaviors, and functional impairment in the former (Magnuson & Constantino, 2011; Oliphant et al., 2020), in line with observations in autistic adults (Newell et al., 2023). Furthermore, autistic youths are overrepresented among patients with treatment-refractory mood disorders (Benarous et al., 2024a) and usually require a longer length of stay in hospitalization compared to other depressed patients (Avrahami et al., 2024; Mandell, 2008).

A major reason for the high rate of untreated depression in autistic individuals is the challenge represented by the assessment of depressive symptoms in this population (Hossain et al., 2020; Kim & Lecavalier, 2021; Lai et al., 2019). Autism features and depressive symptoms can overlap, e.g., social withdrawal, poor facial expressions, and sleep difficulties (Magnuson & Constantino, 2011). An autistic person can find it difficult to understand health care providers' questions about mental states or emotions (e.g., guilt, worthlessness) during clinical interviews (Hinze et al., 2024; Thomas et al., 2024). The diagnosis of depression also requires a comprehensive assessment of symptom duration and change from baseline functioning. This step can require specific adaptations with autistic patients, considering the possible differences for some of them regarding emotional insight or autobiographical memory (Thapar et al., 2023; Westby, 2022). Finally, depressive symptoms in autistic individuals can be mistakenly attributed to co-occurring medical conditions (with pain or physical fatigue) or medication side effects (Magnuson & Constantino, 2011). While the literature on depression in autistic youths is less extensive than for adults, it is expected that most of these barriers are comparable across ages (Kim & Lecavalier, 2021; Magnuson & Constantino, 2011; Thapar et al., 2023).

Limited research exists on the validity and reliability of existing tools for measuring depressive symptoms in autistic individuals. Cassidy et al. (2018) reviewed the studies that provide evidence for the psychometric properties of measurement tools for depressive symptoms in autistic adults. The exclusion by the authors of studies conducted on participants under 18 is regrettable for several reasons. Firstly, while diagnosis criteria for depression in the DSM-5-TR are age-independent, the features of depression can differ across ages, and the specific presentation of pediatric depression may not be adequately captured by current assessment tools developed for adults (Jaffee et al., 2002; Rice et al., 2019). Secondly, most autistic individuals with depression have a first onset of mood symptoms before adulthood, with a first peak of incidence just after puberty in both autistic and non-autistic individuals (Pezzimenti et al., 2019; Thapar et al., 2012). Finally, early recognition and access to appropriate treatment for depression in autistic youths could mitigate the risk of a cascade between negative mood, stressful life experiences, and maladaptive/challenging behaviors, as reported in youths with neurodevelopmental divergences (Benarous et al., 2023; Bernardin et al., 2021; Farhat et al., 2023; Hedley et al., 2018; Rai et al., 2018). A moderate level of evidence supports the efficacy of psychological interventions (e.g., mindfulness and social skill training) to reduce depressive symptoms in autistic individuals (Menezes et al., 2020). In their scoping review, Kim and Lecavalier (2021) identified the lack of information about the psychometric properties of measuring tools for depression in young autistic people as an important gap in the literature and only reported two studies (Mazefsky et al., 2011; Sterling et al., 2015).

This systematic review aims to explore the psychometric properties of the instruments used to assess depressive symptoms in autistic children and adolescents. This research is a first step towards disseminating consensual measurement instruments with satisfactory psychometric properties within the medical community, particularly in the French network of the national expert centers for autism

and complex developmental disorders. A lack of consensual measures contributes to the risk of underestimating depression in this population (Lai et al., 2019; Leyfer et al., 2006). Tools with good psychometric qualities would help to make better decisions regarding the diagnosis and treatment of these young people (Cameron et al., 2021; Pezzimenti et al., 2019; Popow et al., 2021). Given the broad spectrum of autism, the review of the screened studies will pay particular attention to participants' characteristics in terms of demographic (e.g., age, gender) and clinical features (e.g., cognitive level). Specifically, lower intellectual/language skills were linked to additional difficulties in recognizing depressive symptoms in the pediatric population (Benarous et al., 2024b; Hollo et al., 2014), which is especially true for autistic youths (Kim & Lecavalier, 2021).

Methods

The review was conducted using PRISMA 2020 guidelines (Page et al., 2021), and the protocol was registered online with the International Prospective Register of Systematic Reviews (Registration number: CRD42023423377 BLINDED) and can be accessed at (https://www.crd.york.ac.uk/prospero/display_record.php?Record420234233377). Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards are followed in this report (Page et al., 2021).

Search Strategy

The following electronic bibliographic databases were searched on July 2023 using PsycINFO, Embase, Scopus, Science Direct, and PubMed, and data were first extracted on the 3rd of December 2023. An updated database search was conducted in October 2024. The search strategy included the terms shown in Table 1, which were combined using database-specific filters when these were available. The search was restricted to articles in English and those published between January 1980 and the search date. The flow chart (Fig. 1) complies with PRISMA recommendations.

Table 1 General strategy for the review search terms based on Terwee et al. (2009)

Domain	Words
Age group	"children" OR "adolescen*" OR "pediatric" OR "teen" OR "teenager" OR "young people" OR "young person" OR "youth"
Construct search:	"depress*" OR "MDD" OR "mood disorder" OR "affective disorder" OR "dysthym*" OR "melancholy*"
Population search:	"autis*" OR "ASD" OR "ASC" OR "Asperger" OR "Pervasive developmental disorder" OR "PDD-NOS"
Instrument search	"assessment" OR "diagnosis" OR "measure*" OR "questionnaire" OR "psychometr*" OR "interview" OR "screen" OR "scale" OR "checklist" OR "valid*"
Exclusion filter	limited to English language; 1980–November 2024; age 0 to 18 years

Some of these terms were slightly differed according to the electronic bibliographic database

Selection Criteria

One author screened the titles and abstracts of articles (Fig. 1). When there was doubt about whether an article met the inclusion criteria, it was included, and a second author reviewed the full text. Ambiguous papers were discussed with the second reviewer to reach a consensus. The references of the selected articles were hand-searched, and prior recent reviews' reference lists were also found (Cassidy et al., 2018; Hinze et al., 2024; Kim & Lecavalier, 2021; Menezes et al., 2020; Pezzimenti et al., 2019). The following studies were excluded: (1) studies conducted in non-autistic youths, (2) studies conducted in adults, (3) studies where data from pediatric and adult samples were pooled, and (4) studies where depression was represented only as a subscale of a measuring tool for general psychopathology and with no data available on depression. Considering the high rate of overlap within emotional disorders (i.e., anxiety, depression) (Hinze et al., 2024; Pezzimenti et al., 2019), we chose to include instrument tools focusing on emotional disorders in this population.

Data Extraction Method

For each selected study, the following information was noted using a previously tested data extraction form: (i) socio-demographic characteristics (age, gender ratio, countries, type of recruitment); (ii) clinical features (including diagnosis method, Intellectual Quotient (IQ): for estimation of cognitive functioning, and co-occurring psychiatric disorder or neurodevelopmental divergences); (iii) tools used to measure depression, domains captured, and by whom the tool was reported/measured; and (iv) other measures. Once identified, the methodological quality of each article was examined using the Consensus based Standards for the selection of health-based Measurement Instruments (COSMIN) checklist (Mokkink et al., 2010). The checklist considers nine properties of measurement, each with multiple items rated on a 4-point scale: internal consistency, reliability, measurement error, content validity, structural validity, hypothesis testing, criterion validity, responsiveness to change, and cross-cultural validity. Three criteria were used to document the content validity based on the rules provided by Haynes et al. (1995): (a) whether professional experts and a sample of youths with ASD had been involved

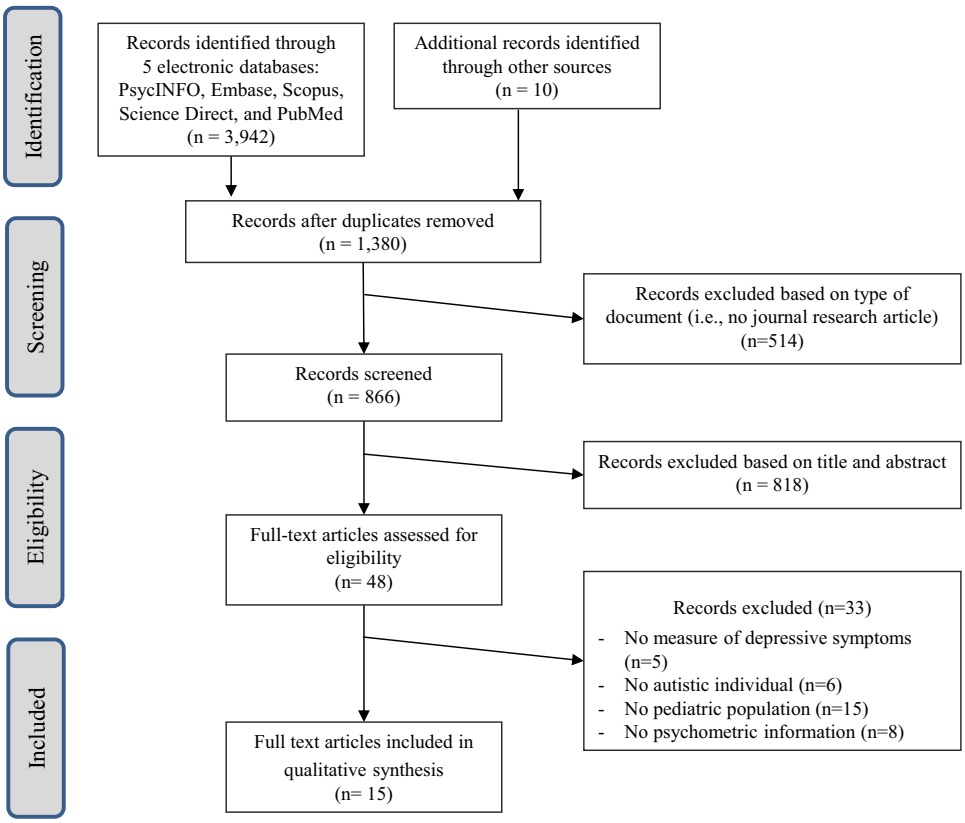

Fig. 1 PRISMA flow chart

in item creation, (b) whether all aspects of the instrument had been submitted to content validity, including form (e.g., modification to item wording, provision of pictorial representations of response scales, inclusion of teaching items) and context of performing, and (c) whether content validity results have been published and updated based on new published data. For each article, the properties addressed are rated excellent, good, fair, or poor, following the guidelines provided by Mokkink et al. (2010). Measurement error was indirectly estimated based on subgroups or correlation analyses testing how the main outcome of the instrument (mean total score, diagnosis) was influenced by participants' socio-demographic or clinical features, as no study provided direct indices such as minimal detectable change (MDC). One reviewer completed the checklists and frequently discussed ratings with a second reviewer to reach a consensus. The second reviewer independently rated all articles using the checklist to check reliability. Agreement on the final rating of each property was 94%, Cohen's κ=.87.

Evidence Synthesis

For each study, a quality rating ranked as positive, indeterminate, or negative was provided for each measurement property examined in line with De Vet et al. (2011). The significance thresholds for validity measurement criteria were based on the European Federation of Psychologists Association' 2013 guideline (EFPA, 2013). For example, internal consistency was positive where Cronbach's alpha is equal to or greater than 0.70; criterion validity was positive where there are convincing arguments that the gold standard is 'gold' and correlation is equal to or greater than 0.70. Finally, the quality ratings for the findings were considered in conjunction with the quality rating for the level of evidence in the articles about each tool.

Results

Description of the Measuring Tools

The search resulted in 866 articles, of which 15 were retained for data extraction (Table 2). In the 15 studies reviewed, a total of twelve different measuring tools were utilized to assess depression in autistic youths. Two of these tools, namely the Autism Comorbidity Interview-Present and Lifetime Version (ACI-PL) (Leyfer et al., 2006) and the Scale for evaluating depressive symptoms among youth on the autism spectrum (EDA) (Bellalou et al., 2021) were specifically developed for individuals with ASD. Six of those tools (i.e., BDI-II, CDI, CDI-S, MFQ-SF, EDA, PHQ-9) measured only depressive symptoms (Bellalou et al., 2021; Bitsika & Sharpley, 2015; Bitsika et al., 2016a; Mazefsky et al., 2011; Santomauro et al., 2016; Uljarević

Table 2 Characteristics of the tools used in included studies

Measure	Rated by	Aim of tool	Nb of items	Subscales	Response options	Population	Used in references	Specific adaptations in NDD population
Questionnaires								
BDI-II (Beck Depression Inventory-II edition)	SR	Assessment of depressive symptoms	21 items	No subscales	4-point Likert scale, ranging from 0 to 3 Total score: 0–63	Adolescents and adults	Original publication: Beck, (1996) ASD: Santomauro, (2016)	No
Child and Adolescent Symptom Inventory– Depression subscale (CASI-D)	SR/IR	Assessment of DSM-IV-TR diagnostics	16 items (anxiety and depressive disorder)	DSM-IV-TR diagnostics	Likert scale 0 to 3 (frequency) or 0/1 (binary)	children and adolescent 6 to 17 years old	ASD: Bitsika, (2016a); Bitsika, (2016b); Bitsika and Sharpley, (2015)	No
Children's Depression Inventory (CDI)	SR	Assessment of depressive symptoms	27 items	No subscales	3-point Likert scale Threshold: 19	Children and adolescents aged 7–17 years	Original publication: Kovacs, (1985) ASD: Mazefsky, (2014); Mazzone et al., (2013)	No
Children's Depression Inventory– Short version (CDI-S)	SR	Assessment of depressive symptoms severity	10 items	No subscales	Likert-type scale 0–2 Total score: 0–54	Children and adolescent ≥7 years old	Original publication: Kovacs, (1992) ASD: Mazefsky et al., (2011)	No
DASS21	SR	Assessment of depression, anxiety and stress	21 items	Depression (7 items), anxiety (7 items), stress (7 items)	4-point Likert scale, ranging from 0 to 3	NS	ASD: Santomauro et al., (2016)	NS
Scale for evaluating depressive symptoms among youth on the autism spectrum (EDA)	IR	Screening for depressive comorbidity of ASD	36 items	pain assessment (13 items), environmental changes (9 items), depressive symptoms (12 items)	Likert-type scale 0- 4	Children and adolescent	Original publication: Bellalou et al., (2021) ASD: Bellalou et al., (2021)	Scale developed and validated in an ASD population
Hospital Anxiety Depression Scale (HADS)	SR	Screening tool for anxiety and depression	14 items	7 items forming the Anxiety subscale—7 items forming the Depression subscale	Likert-type scale 0–3	Children and adolescent	Original publication: Zigmond, (1983) ASD: Uljarevic et al., (2018)	No
Moods and Feelings questionnaire short form (MFQ–SF)	SR	Assessment of depressive symptoms severity	13 items	No	Likert-type scale 0–2 Total score: 0–26 Threshold: 12	Children and adolescent ≥8 years old	Original publication: Angold, (1995) ASD: Andersen et al., (2015); Mazefsky et al., (2014)	No
Patient Health Questionnaire 9 (PHQ-9)	SR	Depression screening in primary care settings (general population)	9 items	No subscale	Likert-type scale 1–5	Children and adolescent	Original publication: Kroenke, (2001) ASD: Pilunthanakul et al., (2021); Uljarevic et al., (2018)	No

Table 2 (continued)

Measure	Rated by	Aim of tool	Nb of items	Subscales	Response options	Population	Used in references	Specific adaptations in NDD population
Revised Children's Anxiety and Depression Scale (RCADS)	SR	Anxiety and depression screening in primary care settings (general population)	47 items	MDD (10 items), PD (9 items), SoP (9 items), SAD (7 items) GAD (6 items), OCD (6 items)	Likert-type scale 0–3	Children and adolescents aged 6–18 years	Original publication: Chorpita, (2005) ASD: Sterling et al. (2015)	No
Semi-structured interviews								
Autism Comorbidity Interview Present and Lifetime version (ACI-PL)	IR	Assessment of depressive symptoms in individuals with ASD (adaptation of K-SADS)	NS	All the subscales present in the K-SADS	4-point scale (0–3)	Children, adolescent and adult	Leyfer et al., (2006), Bellalou et al., (2021) Mazefsky et al., (2011)	Description at the begin of each section of how the manifestation of the disorder in ASD
Kiddie Schedule for Affective Disorders and Schizophrenia Present and Lifetime (KSADS-PL)	Clinician with children and parents	Confirmation of mental disorders in school aged children (coverage of all the psychopathology of children and adolescent)	NS	24 subscales covering of all the psychopathology of children and adolescent	0–3 point rating scale for the majority of items	Children from 6 years old to adolescent until 18 years old	Original publication: Kaufman, (1997) ASD: Pandolfi et al., (2014); Gjevik et al., (2015)	Adaptation of the Schedule for Affective Disorders and Schizophrenia (SADS) for adults developed by Endicott and Spitzer (1978)

SR Self-report, *IR* Informant report

et al., 2018), whereas three of those scales (i.e., HADS, DASS21, RCADS) measured anxiety and depressive symptoms (Santomauro et al., 2016; Sterling et al., 2015; Uljarević et al., 2018). Three scales (CASI, KSADS-PL, ACI-PL) were developed to diagnose depression as part of a global assessment of psychopathology. Seven questionnaires relied solely on self-reports: CDI (Mazefsky et al., 2014) and CDI-S (Mazefsky et al., 2011), HADS (Uljarević et al., 2018), DASS21 (Santomauro et al., 2016), BDI-II (Santomauro et al., 2016), MFQ (Andersen et al., 2015; Mazefsky et al., 2014), PHQ-9 (Pilunthanakul et al., 2021), and RCADS (Sterling et al., 2015), whereas two questionnaires relied solely on informant report: CASI (Bitsika & Sharpley, 2015; Bitsika et al., 2016a, 2016b) and EDA (Bellalou et al., 2021).

The most commonly used measures across these studies were the ACI-PL (used in three studies) and the CASI (Bitsika & Sharpley, 2015; Bitsika et al., 2016b). Four measures were used in two studies each: the KSADS-PL (Gjevik et al., 2015; Pandolfi et al., 2014), the PHQ-9 (Pilunthanakul et al., 2021; Uljarević et al., 2018), the MFQ-SF (Andersen et al., 2015; Mazefsky et al., 2014), and the CDI (Mazefsky et al., 2014; Mazzone et al., 2013). The other measures were used in one study each: the DASS21 (Santomauro et al., 2016), the BDI-II (Santomauro et al., 2016), the CDI-S (Mazefsky et al., 2011), the EDA (Bellalou et al., 2021), the HADS (Uljarević et al., 2018), and the RCADS (Sterling et al., 2015).

Description of the Reviewed Study

The characteristics of the study population for these articles are presented in Table 3. These studies varied in terms of participant demographics, with most of them conducted in the United States ($k=5$) and Australia ($k=4$), Norway ($k=2$), the United Kingdom ($k=1$), Singapore ($k=1$), France ($k=1$) and Italy ($k=1$). The samples included a higher percentage of boys, ranging from 60 to 100%. Most of these studies focused on youths without distinguishing children or adolescents. Age-stratified analyses were not provided in any of

Table 3 Characteristics of study populations in articles on measurement properties

Article	Population/ sample	N	Mean age (SD) years [range]	Male	Country	IQ	ASD Diagnosis	Subgroup analyses	Measures
Leyfer et al. (2006)	Sample from 2 other studies	109	[5–17]	94%	USA	IQ = 82.55 (23.42); [42–141]	ADI-R, ADOS Exclusion of medical cause for ASD	No	ACI-PL (IR)
Mazefsky et al. (2011)	Clinical sample	38	12 (2) [10–17]	82%	USA	IQ > 70 [1]	ADI-R, ADOS	No	CDI-S (SR)
(Mazzone et al., 2013)	Clinical sample	30 ASD 30 MDD 35 HC	ASD: 11.06 (2.59) MDD: 12.76 (2.8) HC: 11.45 (2.3)	100%	Italy	IQ > 85 [1]	ADOS AQ	No	CDI (SR)
Pandolfi et al. (2014)	Clinical sample (outpatient)	76	12 (3)	87%	USA	IQ > 70 [1]	ADI-R, ADOS, clinical consensus	No	KSADS-PL
Mazefsky et al. (2014)	Clinical sample	25 ASD 23 HC	ASD: 15.22 (2.25) HC: 15.56 (2.76)	ASD = 96% HC = 96%	USA	IQ > 80	ADOS ADI-R Expert clinician SRS	No	CDI (SR)
Gjevik et al. (2015)	School Sample	55	11.9 (3.2) [6.3– 17.9]	84%	Norway	NS	ADI-R	No	K-SADS-PL
Bitsika and Sharpley (2015)	School sample	70 ASD 50 HC	2 age groups: [8–12] and [13–18]	100%	Australia	IQ > 90 [1]	DSM-5 ADI-R ADOS	No	CASI-D (SR/IR)
Andersen et al. (2015)	Clinical sample (outpatient)	34 ASD 45 HC	11.6 (2.0) [9–16]	82%	Norway	IQ > 70 [1]	K-SADS-PL ASSQ CBCL	No	MFQ-SF (SR/IR)
Sterling et al. (2015)	Clinical sample	67	12.3 (1.2) [11–15]	NR	USA	IQ > 85 [1]	ADI-R, ADOS	No	RCADS
Santomauro et al. (2016)	Clinical sample	20	15.75 (1.37)	60%	Australia	IQ > 85 [1]	ASDI ASASC	No	BDI-II DASS21
Bitsika et al. (2016b)	Clinical sample (outpatient)	140	11.2 (3.3) [6–18]	100%	Australia	IQ > 70 [1]	DSM-5 ADOS	No	CASI-D (SR/IR)
Bitsika et al. (2016a)	Clinical sample	150 ASD	11.2 (3.3)	100%	Australia	IQ > 70 [1]	ADOS ASDBC	No	CASI-D (SR/IR)
Uljarević et al. (2018)	Clinical samples for two cohorts	151 (106 UK/ 45 Australia)	16.04 (1.28)/ 18.35 (2.55)	71%/ 67%	UK/ Australia	IQ > 70 [1]	SRS, ASQ, clinical diagnosis	Yes (by country)	HADS (SR)
Bellalou et al. (2021)	Clinical sample	153 parents (147 mothers and 6 fathers)	10.71 (3.94) [3–17]	74%	France	IQ not reported (17% wo language)	DSM-5 ADOS, CARS clinical assessment	No	EDA (IR)
Pilunthanakul et al. (2021)	Clinical sample (outpatient)	101	14.6 (2.3)	87%	Singapore	IQ > 70 [1]	DSM or ICD criteria	No	PHQ-9 (SR)

[1]Eligibility criteria

ACI-PL Autism Comorbidity Interview- Present and Lifetime Version, *ADI-R* Autism Diagnostic Interview Revised, *ADIS-C/P* Anxiety Disorders Interview Schedule-Child/Parent, *ADOS* Autism Diagnostic Observation Scheduled, *AQ* Autism Questionnaire, *ASASC* Australian Scale for Autism Spectrum Conditions, *ASD* Autism Spectrum Disorder, *ASDBC* Autism Spectrum Disorder Behaviour Checklist, *ASDI* Asperger Syndrome (and High-Functioning Autism) Diagnostic Interview, *ASQ* Autism Spectrum Questionnaire, *ASSQ* Autism Spectrum Screening Questionnaire, *BDI* Beck Depression Inventory, *CARS* Childhood Autism Rating Scale, *CASI* Child and Adolescent Symptom Inventory, *CASS-S* Conners-Wells Adolescent Self-Report Scale-Short Edition, *CBCL* Child Behavior Checklist, *CDI* Children's Depression Inventory, *DASS21* Depression Anxiety Stress Scale 21 items, *EDAEchelle d'évaluation de la symptomatologie dépressive spécifique à l'autisme* or Scale for evaluating depressive symptoms among youth on the autism spectrum, *HADS* Hospital Anxiety Depression Scale, *HC* Healthy Controls, *IR* Informant-Report (parents in most cases), *KSADS-PL* Kiddie Schedule for Affective Disorders and Schizophrenia Present and Lifetime, *MDD* Major Depressive Disorders, *MFQ-SF* Mood and Feeling Questionnaire Short Form, *NR* Not Reported, *PHQ-9* Patient Health Questionnaire 9, *SR* Self-Report, *SRS* Social Responsiveness Scale

the studies. Most articles excluded individuals with intellectual development disorder (IDD) from their recruitment criteria (Bitsika & Sharpley, 2015; Bitsika et al., 2016a; Leyfer et al., 2006; Mazefsky et al., 2011, 2014; Mazzone et al., 2013; Pandolfi et al., 2014; Pilunthanakul et al., 2021; Santomauro et al., 2016; Sterling et al., 2015; Uljarević et al., 2018). Regarding sample size, half of the studies included less than a hundred participants (Andersen et al., 2015; Bellalou et al., 2021; Bitsika et al., 2016a, 2016b; Leyfer et al., 2006; Mazefsky et al., 2011, 2014; Mazzone et al., 2013; Pandolfi et al., 2014; Pilunthanakul et al., 2021; Santomauro et al., 2016; Sterling et al., 2015; Uljarević et al., 2018). Twelve studies out of fourteen were conducted on clinical samples (Andersen et al., 2015; Bellalou et al., 2021; Bitsika et al., 2016a, 2016b; Leyfer et al., 2006; Mazefsky et al., 2011, 2014; Mazzone et al., 2013; Pandolfi et al., 2014; Pilunthanakul et al., 2021; Santomauro et al., 2016; Sterling et al., 2015; Uljarević et al., 2018), and two on school samples (Bitsika & Sharpley, 2015; Gjevik et al., 2015).

Psychometric Properties of Measuring Tools

Table 4 presents each article's methodological quality. The most documented psychometric properties were internal consistency ($k=12$), criterion validity ($k=10$), and measurement errors ($k=6$).

The internal consistency of the scales across instruments was mainly good/excellent, with a Cronbach's alpha (α) correlation for all items above 0.7 for the CDI and CDI-S (Mazefsky et al., 2011), the CASI (Bitsika & Sharpley, 2015; Bitsika et al., 2016a, 2016b), and above 0.8 for the EDA (Bellalou et al., 2021), the PHQ-9 (Pilunthanakul et al., 2021), the MFQ-SF (Mazefsky et al., 2014), the DASS21 (Santomauro et al., 2016), the BDI-II (Santomauro et al., 2016) and the RCADS (Sterling et al., 2015). For the KSADS-PL (Pandolfi et al., 2014), the authors reported Guttman's Lambda-2 (λ) as an alternative reliability estimate for Cronbach's alpha.

Evidence of inter-rater and test–retest reliability was found for the ACI-PL and the PHQ-9. The inter-rater reliability was estimated to be excellent for the ACI-PL using the diagnosis criteria for depression reformulated from the original KSADS-PL (inter-rater reliability $\kappa=0.8$, test-rest reliability for lifetime depression $\kappa=0.6$) (Leyfer et al., 2006) and poor for the PHQ-9 with an optimal cutoff at 5 based on ROC analysis (inter-rater reliability for PHQ-9 total score, $r=.49$) (Pilunthanakul et al., 2021).

The measurement error was good/fair for the CDI (Mazefsky et al., 2011), the EDA (Bellalou et al., 2021), the HADS (Uljarević et al., 2018), the PHQ-9 (Pilunthanakul et al., 2021), and the CASI (Bitsika & Sharpley, 2015; Bitsika et al., 2016b). No study provided direct index of measurement error.

The criterion validity was excellent for the ACI-PL (sensitivity 100%, specificity 83–94% to identify cases receiving treatment for depression) (Leyfer et al., 2006), good for the EDA (based on correlation tests comparing scores between participants and an external evaluator in a sub-sample of 14 participants) (Bellalou et al., 2021), good and fair for the KSADS-PL (sensitivity 100%, specificity 82% to identify those scoring high to CBCL affective problems subscale for Gjevik et al. (2015)) and correlations between KSADS severity of the depression and CBCL Anxious/Depressed scale $r=.68$, withdrawn depressed scale $r=.55$ for Pandolfi et al. (2014)). The criterion validity was fair for the HADS (based on correlations with other depression scales) (Uljarević et al., 2018), the PHQ-9 (based on diagnosed cases of depression using the KSADS-PL) (Pilunthanakul et al., 2021), and the RCADS (based on correlations with CBCL) (Sterling et al., 2015), and fair and poor for the CDI (sensitivity 27%, specificity 89% to identify cases diagnosed with the ACI-PL for Mazefsky et al. (2011) and significant relation between CDI total score and CBCL internalizing subscale for Mazzone et al. (2013), and poor for the MFQ-SF (no significant correlation with others scales in Andersen et al. (2015) and significant correlation only with one of the two scales for internalizing symptoms in Mazefsky et al. (2014)).

The structural validity was explored for the EDA (Bellalou et al., 2021), the HADS (Uljarević et al., 2018), and the CASI (Bitsika & Sharpley, 2015; Bitsika et al., 2016b). The EDA is composed of two factors: (1) behavioral changes and (2) emotional and cognitive changes, which were explored by two different EFA (exploratory factor analysis) with Kaiser–Meyer–Olkin index greater than 0.80 (KMO1 = 0.91 and KMO2 = 0.92). The internal consistency of the first factor was excellent ($\alpha F1=0.91$; $\omega F1=0.91$) and very good for the second factor ($\alpha F2=0.83$; $\omega F2=0.82$). The HADS structural validity was explored using principal component analysis (PCA) in two samples: one in the UK and one in Australia (Uljarević et al., 2018). The PCA indicated two factors as the optimal solutions in both samples. The KMO was 0.78 for the UK and 0.74 for the Australian sample. Reliability for the whole sample internal consistency (Cronbach's α) was good $\alpha=.83$ (UK: $\alpha=.84$; Australian: $\alpha=.82$) for the HADS-A scale and acceptable $\alpha=.65$ (UK: $\alpha=.60$; Australian: $\alpha=.73$) for HADS-D scale.

Overall Assessment of Synthesized Evidence

Table 5 presents the synthesized evidence on the quality of the measurement properties of the reviewed tools.

The EDA is the only measure with five domains of psychometric properties (internal consistency, measurement

Table 4 Methodological quality of each article per measurement property and instrument according to COSMIN Checklist

Measure	Article	Version	Internal consistency	Inter-raters and test–retest reliability	Measurement error	Content validity	Structural validity	Hypothesis testing	Criterion validity	Responsiveness to change	Cross-cultural validity
Questionnaires											
BDI-II	Santomauro et al., 2016	SR	Excellent *BDI-II total score α=.94*	na	na	na	na	na	na	Poor *No significant change in time after intervention*	na
CASI-D	Bitsika et al., 2016b[1]	SR/IR	Good *α=.79 for SR α=.86 for IR*	na	Good *No significant difference based on age and IQ scores*	na	Good *Factorial analysis identifies four factors (KMO>0.6 for SR/IR)*	Good *Correlations between some factors and biomarkers of stress (salivary cortisol)*	na	na	na
	Bitskia et al., 2016a[2]	SR/IR	Good *α=.75 for SR α=.75 for IR*	na	na	na	Good *Factorial analysis identifies three factors (KMO>0.6 for SR/IR)*	na	na	na	na
	Bitskia and Sharpley, 2015[2]	IR/SR	Good *α=.70*	na	Good *No significant difference based on IQ level*	na	na	na	na	na	na
CDI	Mazzone et al., 2013	SR	na	na	na	na	na	na	Fair *Significant relation between CDI total score and CBCL internalizing subscale (R^2=.29)*	na	na
	Mazefsky et al., 2011	SR	Good *α=.80*	na	Fair *No difference based on IQ level*	na	na	na	Poor *Compared to DSM depression diagnosis with ACI-PL, sensitivity 27%, specificity 89%*	na	na
DASS21	Santomauro et al., 2016	SR	Excellent *DASS depression subscale α=.94*	na	na	na	na	na	na	Good *No significant change in pre/post-intervention*	na

Table 4 (continued)

Measure	Article	Version	Internal consistency	Inter-raters and test–retest reliability	Measurement error	Content validity	Structural validity	Hypothesis testing	Criterion validity	Responsiveness to change	Cross-cultural validity
EDA	Bellalou et al., 2021	IR	Excellent *α=.92* *ω=.92*	na	Good *No difference based on language level*	Excellent *Preliminary study to assure understandability and review by group of experts*	Excellent *Factorial analysis found three factors* *KMO*=0.9	na	Good *Correlations diagnosis of depression by a psychiatrist in a subsample (n=14) (r=.73)*	na	na
HADS	Uljarevic et al., 2018	Self-report	Moderate *For HADS depression* *α=.65*	na	Fair *No effect of age*	na	Good *Factorial analysis identifies two factors (KMO=0.82)*	na	Fair *Correlations between HADS-Depression and PHQ-9 depression (r=.56) [only in one out of two sample], and also HADS-Anxiety, SDQL emotional scale and negative correlation with mental well being*	na	Good *No significant difference between UK and Australian samples*
MFQ-SF	Andersen et al., 2015	SR/IR	na	na	na	na	na	Poor *No correlation between change in MFQ-Short Form and change in working memory, or cognitive flexibility*	Poor *No correlation between the MFQ-SF and the CBCL affective scales at baseline, follow-up*	Good *Change in depressive symptoms over time*	na
	Mazefsky et al., 2014	SR/IR	Excellent *α=.85 for SR* *α=.86 for IR*	na	na	na	na	Good *Correlation of SR/IR with some maladaptive emotional regulation strategies*	Poor *Correlation not significant between MFQ-SF and CBCL internalizing scale (r=-.16) but significant with YSR internalizing (r=.69)*	na	na

Table 4 (continued)

Measure	Article	Version	Internal consistency	Inter-raters and test–retest reliability	Measurement error	Content validity	Structural validity	Hypothesis testing	Criterion validity	Responsiveness to change	Cross-cultural validity
PHQ-9	Pilunthanakul et al., 2021	SR/IR	Excellent *α=.81 for SR* *α=.85 for IR*	Fair *IRA, r=.49*	Fair *No effect of age or education*	na	na	na	Fair *Correlation with depression diagnosis with MINI-Kid,* *AUC (SR)=.65 AUC (PR)=.72*	na	na
RCADS	Sterling et al., 2015	SR	Excellent *α=.72 to.93*	na	na	na	na	na	Fair *Significant relation between RCADS internalizing score and CBCL Anxiety/depression (r=.31) but no other CBCL subscales*	na	na
Semi-structured interviews											
ACI-PL	Leyfer et al., 2006	SR/IR	na	Excellent *IRA: 90%, κ=0.8* *TRR for lifetime diagnosis of major depression: κ=0.6*	na	Excellent *Detailed description of adaptations based on reports in the literature and clinical experience*	na	na	Excellent *All identified case received treatment for major depression sensitivity 100%* *specificity 83–94%*	na	na

Table 4 (continued)

Measure	Article	Version	Internal consistency	Inter-raters and test–retest reliability	Measurement error	Content validity	Structural validity	Hypothesis testing	Criterion validity	Responsiveness to change	Cross-cultural validity
KSADS-PL	Pandolfi et al., 2014	SR/IR	Good *For KSADS depression module λ-2=.79*	na	na	na	na	na	Fair *Moderate correlation between KSADS depression scores and CBCL Anxious/Depressed scale (r=.68), Withdrawn depressed scale (r=.55)*	na	na
	Gjevik et al., 2015	SR/IR	na	na	na	na	na	na	Good *Correlation between CBCL Affective Problems versus any depressive disorder (k=0.5, sensitivity 100%, specificity 82%)*	na	na

ACI-PL Autism Comorbidity Interview- Present and Lifetime Version, *ASASC* Australian Scale for Autism Spectrum Conditions, *ASDBC* Autism Spectrum Disorder Behaviour Checklist, *CAPA* Child and Adolescent Psychiatric Assessment-parent version, *CASI-D* Child and Adolescent Symptom Inventory-Depressive symptoms subscale, *CBCL 6/18* Child Behavior Checklist 6/18, *CDI-S*: Children's Depression Inventory Short version, *CASS-S* Conners-Wells Adolescent Self-Report Scale-Short Edition; *DASS* Depression Anxiety Stress Scal, *EDA* Echelle d'évaluation de la symptomatologie dépressive spécifique à l'autisme or Scale for evaluating depressive symptoms among youth on the autism spectrum, *ERQ* Emotion Regulation Questionnaire, *HADS* Hospital Anxiety Depression Scale, *HMHH-ATN* Health and Mental Health History Form–Autism Treatment Network, *KSADS-PL* Kiddie Schedule for Affective Disorders and Schizophrenia; *PHQ-9* Patient Health Questionnaire 9, *RCMAS* Revised Children's Manifest Anxiety Scale, *SLOI-CV* Short Leyton Obsessional Inventory-Child Version (SLOI-CV), *SPr* Sensory Profile, *SR* Self-Report, *RMET* Reading the Minds in the Eyes Task, *YSR* Youth Self Report, *IR* Informant-Report (parents in most case)

Table 5 Quality of measurement properties of tools

Measure	Version	Measurement properties								
		Internal consistency	Reliability	Measurement error	Content validity	Structural validity	Hypothesis testing	Criterion validity	Responsiveness to change	Cross-cultural validity
Questionnaires										
BDI-II	SR	+	na	na	na	na	na	na	-	na
CASI-D	IR/SR	++	na	++	na	++	+	na	na	na
CDI	SR	+	na	±	na	na	na	±	na	na
DASS21	SR	+	na	na	na	na	na	na	+	na
EDA	IR	+++	na	++	+++	+++	na	++	na	na
HADS	SR	++	na	+	na	++	na	+	na	++
MFQ-SF	IR/SR	++	na	na	na	na	±	–	+	na
PHQ-9	SR	+++	+	+	na	na	na	+	na	na
RCADS	SR	+	na	na	na	na	na	+	na	na
Semi-structured interviews										
ACI-PL	SSI	na	+++	na	+++	na	na	+++	na	na
KSADS-PL	SSI	++	na	na	na	na	na	++	na	na

SR self-report, *IR* informant report. Scoring system: Strong evidence for (+++) or against (–) a measurement property/Moderate evidence for (++) or against (–) a measurement property/Limited level of evidence for one study of fair quality (+ or −) Conflicting evidence (±)/Unknown (?) evidence/na = no information available

error, content validity, structural validity, and criterion validity) that is assessed as having strong or moderate evidence for (++/+++). No evidence supports the inter-rater or test–retest reliability, responsiveness to change, hypothesis testing, and cross-cultural validity for this scale.

Strong evidence supports the internal consistency of the PHQ-9 in youths with ASD, with limited evidence for reliability, measurement error, and criterion validity. Moderate evidence supports the internal consistency, structural validity, and cross-cultural validity of the HADS, with limited evidence for measurement error and criterion validity. The review found moderate evidence for the CASI-D's internal consistency, measurement error, and structural validity, as well as limited evidence for hypothesis testing.

Very little evidence was available regarding the psychometric properties of other questionnaires (i.e., BDI-II, CDI, CDI-S, MFQ-SF, RCADS, and DASS21). The evidence for internal consistency was consistently reported as limited for these four questionnaires.

The ACI-PL has three domains of psychometric properties (reliability, content validity, and criterion validity), which are assessed as having strong or moderate evidence for (++/+++). The KSADS-PL had moderate evidence supporting internal consistency and criterion validity.

Discussion

Summary of the Main Findings

The review identified 15 empirical studies that examined the properties of twelve different measurement tools for assessing depression in autistic youths. Only two instruments were developed explicitly for autism (i.e., EDA and ACI-PL). Eight questionnaires developed for non-autistic youths were used without significant adaptation (i.e., HADS, PHQ-9, CASI, BDI-II, CDI, CDI-S, MFQ-SF, DASS21, and RCADS). One semi-structured interview was used without significant adaptation (i.e., KSADS-PL).

Among the identified tools, the EDA tool stands out for its robust psychometric qualities and high level of reliability across various areas studied. Its development is well-detailed, including a preliminary step to document content validity (Bellalou et al., 2021). The scale integrates additional features of depression specifically described in autistic youths, such as aggressive behaviors or increased stereotypies (Magnuson & Constantino, 2011). In addition, the EDA was, with the ACI-PL, the only instrument tested in autistic individuals with communication support needs. However, its length poses a challenge for its use in clinical routine, as it requires documenting fifteen domains of a patient's functioning with parents or other caregivers, making it more time-consuming than the other tools. While these findings are encouraging, they need replication by different research teams and in distinct samples. Further research is also required to determine the extent to which the EDA can

detect clinically significant changes in depressive symptoms, for example, after a therapeutic intervention.

Three questionnaires have evidence supporting intermediary psychometric qualities: the HADS, the PHQ-9, and the CASI. The HADS is a 14-item self-report questionnaire initially developed to identify both anxiety and depressive symptoms. It was designed to screen out individuals with somatic illness, so it does not address 'biological' symptoms of depression (i.e., sleep or appetite problems) despite their importance in documenting depression in patients with communication support needs and/or neurodevelopmental divergences (Magnuson & Constantino, 2011). The HADS items mainly focus on cognitive symptoms of depression, and most formulations concern the patient's subjective experience and then require some degree of insight (e.g., item 2, '*I still enjoy the things I used to enjoy*'). Considering the complexity of the item formulation, its validity can be questioned for autistic individuals with specific communication support needs or with co-occurring IDD (Magnuson & Constantino, 2011). The PHQ-9, whose psychometric qualities are generally comparable, has the advantage of being relatively short (9 items only) and provides a more straightforward formulation of items, focusing on behavior rather than cognition (e.g., item 7, '*Trouble concentrating on things*'—note the lack of personal pronoun in the sentence). Finally, the CASI, used in three studies conducted by the same research team, consists of items formulated on the model of the DSM criteria.

The other questionnaires (i.e., the BDI-II, the CDI, the MFQ-SF, RCADS, and the DASS21) showed poor psychometric qualities. One of the main limitations of using these questionnaires on autistic youths can be the length of statements (21 items for the BDI-II, 27 items for the CDI, 21 items for the DASS21, 47 items for the RCADS), sometimes complex or even involving double negatives (e.g., item 11 of the CDI). The MFQ-SF, which has a simpler structure of 13 items, encompasses symptoms relating to stress and emotional dysregulation, not specific to depression.

The studies using semi-structured interviews to diagnose depressive disorders in autistic youths (i.e., KSADS-PL and ACI-PL) have shown positive results regarding the psychometric quality of these tools. The ACI-PL was derived from the KSADS-PL to better match the particularities of ASD patients by developing items specifically related to depression. As the domains of the psychometric qualities studied were distinct for the KSADS-PL and the ACI-PL, the two instruments cannot be directly compared. Then, the benefit of using specific rather than generic semi-structured interviews cannot be determined from the reviewed findings.

Study Limitations

The following limitations are worth considering when interpreting the findings. The sample sizes in many of these studies were relatively small, which may limit the generalizability of the results. Additionally, the inclusion criteria sometimes included co-occurring neurodevelopmental divergences, which may introduce selection bias and affect the external validity. Such selection bias may be a source of particular concern as such co-occurrence is precisely recognized as a risk factor for a more resistant form of depression in autistic individuals (O'Connor et al., 2023) and an additional challenge for a correct diagnosis of mood disorder (Magnuson & Constantino, 2011). Considering the lack of subgroup analysis, the confounding effects of co-occurring neurodevelopmental divergences or mental health issues cannot be adequately studied. Such limitation is also reported in other systematic reviews on comparable topics (Benarous et al., 2024b). Consensual recommendations, such as the COSMIN checklist and the European Federation of Psychologists Association's thresholds, were used to limit heterogeneity in interpreting the results of varied statistical analyses. The inclusion of studies mainly conducted in males and with a wide age range can also be regarded as a possible limitation on the generalizability of findings. Finally, publication bias may exist as articles written in a language other than English may have been omitted.

Research and Clinical Implications

While there is extensive research exploring the co-occurrence of neurodevelopmental divergences and psychiatric disorders in autistic individuals (Hossain et al., 2020), the identification and treatment of pediatric depression remains challenging (Kim & Lecavalier, 2021; Menezes et al., 2020; Pezzimenti et al., 2019) despite its presumed high prevalence in this population (Hudson et al., 2019; Lai et al., 2019; Leyfer et al., 2006). It is recommended to use standardized assessment tools to evaluate associated psychopathology in autistic children and adolescents (Volkmar et al., 2014). Of course, as for non-autistic youths, good diagnosis practice requires using multi-method multi-informants to assess symptoms in multiple contexts, considering the limitations inherent in each data collection method (Hinze et al., 2024; Pezzimenti et al., 2019). However, one limitation currently faced is the lack of consensus regarding a definitive measure for capturing depression, specifically in autistic individuals (Magnuson & Constantino, 2011). In this population, the onset of depressive symptoms during childhood or adolescence can partly reflect the challenges that these individuals can experience in addressing multiple transitions at the individual, family, and social levels while their specific needs

for communication or sensory-integration issues are insufficiently addressed (Thapar et al., 2023). Two longitudinal studies showed that school functioning and poor social support were important mediators in the relations between the severity of autistic traits and the onset of depression in late adolescence (Hedley et al., 2018; Rai et al., 2018).

Although several tools have been developed for use within this population, this review showed that only a limited number of findings support the quality of these instruments. With the limitation mentioned above, some of these tools can provide valuable information about the presence and severity of depressive symptoms in this population, in combination with other sources of information. Considering the total number of reviewed studies, the high number of instruments indicates no consensus among researchers, with low evidence for each of them. While the EDA was the instrument whose psychometric properties were the most well-investigated, more is needed to confirm its validity and discuss its feasibility in clinical routine. It is also essential for future studies to provide more information on the "external" and not only the "internal" validity of these tools. Such studies could examine the sensitivity and specificity of these tools in detecting depressive symptoms in autistic youths in comparison with symptoms of other psychiatric disorders and neurodevelopmental divergences (e.g., anxiety, attention deficit disorder) or symptoms related to factors commonly observed in this population (e.g., preexisting specificities of communication, specialized interests, or pain). Besides, it is worth reminding that these instruments were developed and tested for different purposes. Some questionnaires are helpful for screening for depression (e.g., the HADS). Other instruments are useful for validating a diagnosis following an interview (e.g., KSADS-PL). Finally, other measures are helpful to catch the change in the severity of depressive symptoms after an intervention.

Based on the reviewed findings, the PHQ-9 or the HADS should be the first option when clinicians want to use a self-report instrument. Interestingly, in the systematic review conducted by Cassidy et al. (2018), the PHQ-9 was also regarded as one of the most robust tools for measuring depression in autistic adults. Interpreting the patient's scores (and eventually subscores) to other reviewed self-report questionnaires is difficult considering the imprecision about the measured clinical phenomenon. For example, the vague formulation of the MFQ-SF items raises questions about its content validity and its ability to discriminate between depressive disorders vs. chronic emotional dysregulation. Considering the possible inaccuracies and biases with both self-report (e.g., influenced by factors such as cognitive and communication competencies) and informant-report measures in assessing depressive symptoms in autistic individuals (Thomas et al., 2024), the use of combined sources of information (i.e., young, family, and other caregivers) should be privileged (Lai et al., 2019). One could also regret that no study provides information about the benefit of using combined approaches versus only one source of information (self- or informant-report).

It is striking to note that only two studies included participants with an IQ below 70 (Bellalou et al., 2021; Leyfer et al., 2006), and none included specifically a group of participants with IDD. While the poor representativeness of individuals with IDD is a recurrent problem in autism research (Thurm et al., 2022), this is particularly problematic here as cognitive deficit may influence the prevalence rates of depressive disorders in autistic individuals. The meta-analysis conducted by Hudson et al. (2019) found that the greater mean IQ of the sample was associated with greater lifetime prevalence, while the percentage of IDD was not found to be a significant moderator in the meta-analysis conducted by Lai et al. (2019). Some authors suggested that higher cognitive skills are associated with better insight into their limitations, resulting in increased depression risk (Chandrasekhar & Sikich, 2015). In contrast, others stressed the higher likelihood of medical (e.g., pain, co-occurring physical illness), psychological (e.g., lower problem-solving capacity), and social risk factors (e.g., adverse life events, lack of intimate relationships) for depression in autistic youths with IDD compared to those without (Mazurek & Kanne, 2010). The lack of valid and consensual instruments for measuring depression in autistic youths with language/cognitive deficits could partly explain the relative paucity of studies in this population.

A significant overrepresentation of male participants was noted among the reviewed studies (Table 3), with an average of 87% male participants, which significantly exceeds the expected gender ratio for ASD of approximately 1:4. This discrepancy is seen as problematic, as the proportion of females has been linked to a higher prevalence of depressive disorders in the studies reviewed in the meta-analysis conducted by Lai et al. (2019), aligning with the epidemiology of depression observed in neurotypical individuals (Thapar et al., 2012). The extent to which gender differences are present in pediatric samples varies across ages and involves gender-specific risk factors (e.g., biological influences, adverse life events) remains to be further explored. Notably, based on longitudinal data from 165 participants (n = 109 with ASD; n = 56 with IDD without ASD), Gotham et al. (2015) found an increase in depressive symptoms throughout adolescence among autistic girls, while males showed a more consistent trajectory of depressive symptoms from childhood through young adulthood. The higher burden of depression in autistic girls relative to boys is further supported by studies indicating a greater incidence of

suicidality in autistic females compared to their male counterparts (Hirvikoski et al., 2020; Newell et al., 2023).

Finally, if further research would certainly contribute to the development of more appropriate tools to help clinicians diagnose depression in autistic youths, other barriers to appropriate treatments are worth considering. Firstly, aligning with the dimensional model of autism in the DSM-5-TR (American Psychiatric Association, 2022), clinicians may erroneously interpret all patients' emotional and behavioral problems as directly related to the autism spectrum. The risk of adopting such an exclusive view is to mask potentially treatable co-occurring psychiatric disorders or medical conditions (Astle et al., 2022; Hinze et al., 2024). One exception should be mentioned with catatonia, which has been included as a specifier that poses challenges in diagnosis (Hauptman et al., 2023) and treatment (Guinchat et al., 2020). Secondly, mental health professionals would certainly be more prone to screen out depression if they are confident that a positive case would have direct implications for the patient's treatment plan. Much could be done to provide adapted training to mental health care providers in how therapeutics usually recommended for pediatric depression (i.e., behavioral therapy or antidepressant) can be successfully adapted for autistic patients (Benarous et al., 2024a; Cameron et al., 2021; Menezes et al., 2020; Popow et al., 2021; Santomauro et al., 2016).

Conclusions

Our review finds 15 empirical studies presenting the psychometric properties of twelve tools for assessing depression in autistic youths. Among measuring tools for evaluating depression in autistic youths, the EDA stands out as having the most reliable and valid psychometric properties. The empirical evidence supporting the quality of questionnaires not explicitly designed to be used in the population with developmental disabilities was limited. However, the PHQ-9 can be helpful for screening for depression using a self-report questionnaire.

Acknowledgments No specific information.

Author Contributions Study concept and design: XB, JMG, RL, HL, HY. Acquisition of data: XB, HL, HY. Interpretation of data: XB, JMG, RL, HL, DC, CC, AC, HY. Drafting the manuscript: XB, JMG, RL, HL, DC. Critical revision of the manuscript for important intellectual content: JMG, RL, HL, DC, CC, AC, HY. Final draft: All authors.

Funding No research grant supports this project.

Declarations

Competing Interests All coauthors certify that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

References

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association.

Andersen, P. N., Skogli, E. W., Hovik, K. T., Egeland, J., & Øie, M. (2015). Associations among symptoms of autism, symptoms of depression and executive functions in children with high-functioning autism: A 2 year follow-up study. *Journal of Autism and Developmental Disorders, 45*, 2497–2507. https://doi.org/10.1007/s10803-015-2415-8

Angold, A., Costello, E. J., Messer, S. C., & Pickles, A. (1995). Development of a shortquestionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*, *5*(4), 237–249

Astle, D. E., Holmes, J., Kievit, R., & Gathercole, S. E. (2022). Annual research review: The transdiagnostic revolution in neurodevelopmental disorders. *Journal of Child Psychology and Psychiatry, 63*, 397–417. https://doi.org/10.1111/jcpp.13481

Avrahami, M., Ben-Dor, D. H., Ratzon, R., Weizman, A., & Perlman Danieli, P. (2024). Characterizing the clinical and sociodemographic profiles of hospitalized adolescents with autism spectrum disorder. *Global Mental Health (Cambridge Prisms), 11*, e63. https://doi.org/10.1017/gmh.2024.63

Beck, A. T., Steer, R. A., & Brown, G. (1996). Beck Depression Inventory–II (BDI-II)[Database record]. *APA PsycTests*. https://doi.org/10.1037/t00742-000

Bellalou, L., Downes, N., & Cappe, E. (2021). Development and preliminary validation of a depressive symptomatology detection scale among children and adolescents on the autism spectrum. *Autism, 25*, 361–373. https://doi.org/10.1177/1362361320958209

Benarous, X., Guilé, J. M., & Consoli, A. (2023). Letter to the editors: An interactive model to understand the vulnerability to emotional disorders in youths with neurodevelopmental disorders-a triple jeopardy? *European Child and Adolescent Psychiatry*. https://doi.org/10.1007/s00787-023-02212-w

Benarous, X., Lefebvre, C., Guilé, J. M., Consoli, A., Cravero, C., Cohen, D., & Lahaye, H. (2024). Effects of neurodevelopmental disorders on the clinical presentations and therapeutic outcomes of children and adolescents with severe mood disorders: A multicenter observational study. *Child Psychiatry & Human Development*. https://doi.org/10.1007/s10578-024-01770-7

Benarous, X., Walesa, S., Guilé, J. M., Cravero, C., Consoli, A., Cohen, D., Young, H., Labelle, R., & Lahaye, H. (2024). A systematic review of the psychometric properties of tools for measuring depression in youths with intellectual disability. *European Child & Adolescent Psychiatry*. https://doi.org/10.1007/s00787-024-02405-x

Bernardin, C. J., Lewis, T., Bell, D., & Kanne, S. (2021). Associations between social camouflaging and internalizing symptoms in autistic and non-autistic adolescents. *Autism, 25*, 1580–1591. https://doi.org/10.1177/1362361321997284

Bitsika, V., & Sharpley, C. F. (2015). Differences in the prevalence, severity and symptom profiles of depression in boys and adolescents with an autism spectrum disorder versus normally developing controls. *International Journal of Disability, Development*

and Education, 62, 158–167. https://doi.org/10.1080/1034912X.2014.998179

Bitsika, V., Sharpley, C. F., Andronicos, N. M., & Agnew, L. L. (2016a). Prevalence, structure and correlates of anxiety-depression in boys with an autism spectrum disorder. *Research in Developmental Disabilities, 49–50*, 302–311. https://doi.org/10.1016/j.ridd.2015.11.011

Bitsika, V., Sharpley, C. F., & Mills, R. (2016b). Are sensory processing features associated with depressive symptoms in boys with an ASD? *Journal of Autism and Developmental Disorders, 46*, 242–252. https://doi.org/10.1007/s10803-015-2569-4

Cameron, L. A., Phillips, K., Melvin, G. A., Hastings, R. P., & Gray, K. M. (2021). Psychological interventions for depression in children and young people with an intellectual disability and/or autism: Systematic review. *British Journal of Psychiatry, 218*, 305–314. https://doi.org/10.1192/bjp.2020.226

Cassidy, S. A., Bradley, L., Bowen, E., Wigham, S., & Rodgers, J. (2018). Measurement properties of tools used to assess depression in adults with and without autism spectrum conditions: A systematic review. *Autism Research, 11*, 738–754. https://doi.org/10.1002/aur.1922

Chandrasekhar, T., & Sikich, L. (2015). Challenges in the diagnosis and treatment of depression in autism spectrum disorders across the lifespan. *Dialogues in Clinical Neuroscience, 17*, 219–227. https://doi.org/10.31887/DCNS.2015.17.2/tchandrasekhar

Chorpita BF, Moffitt CE, Gray J. (2005) Psychometric properties of the Revised Child Anxiety andDepression Scale in a clinical sample. *Behav Res Ther.*, *43*(3):309−22. https://doi.org/10.1016/j.brat.2004.02.004. PMID: 15680928

De Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine*. Cambridge University Press.

EFPA (2013). European Federation of Psychologists' Association review model for the description and evaluation of psychological tests: Test review form and notes for reviewers.

Endicott, J., & Spitzer, R. L. (1978). A diagnostic interview: The schedule for affective disorders and schizophrenia. *Archives of General Psychiatry*, *35*(7), 837–844. https://doi.org/10.1001/archpsyc.1978.01770310043002

Farhat, L. C., Blakey, R., Davey Smith, G., Fujita, A., Shephard, E., Stergiakouli, E., Eley, T. C., Thapar, A., & Polanczyk, G. V. (2023). Networks of neurodevelopmental traits, socioenvironmental factors, emotional dysregulation in childhood, and depressive symptoms across development in two UK cohorts. *American Journal of Psychiatry, 180*(10), 755–765.

Gjevik, E., Sandstad, B., Andreassen, O. A., Myhre, A. M., & Sponheim, E. (2015). Exploring the agreement between questionnaire information and DSM-IV diagnoses of comorbid psychopathology in children with autism spectrum disorders. *Autism, 19*, 433–442. https://doi.org/10.1177/1362361314526003

Gotham, K., Brunwasser, S. M., & Lord, C. (2015). Depressive and anxiety symptom trajectories from school age through young adulthood in samples with autism spectrum disorder and developmental delay. *Journal of the American Academy of Child and Adolescent Psychiatry, 54*, 369–76.e3. https://doi.org/10.1016/j.jaac.2015.02.005

Guinchat, V., Cravero, C., Lefèvre-Utile, J., & Cohen, D. (2020). Multidisciplinary treatment plan for challenging behaviors in neurodevelopmental disorders. *Handbook of Clinical Neurology, 174*, 301–321. https://doi.org/10.1016/b978-0-444-64148-9.00022-3

Gussin, H. A., Shiu, C. S., Danguilan, C., Mihaila, I., Acharya, K., & Berg, K. L. (2024). Impact of adverse childhood experiences and mental health on school success in autistic children: Findings from the 2016–2021 national survey of children's health. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-024-06338-x

Hauptman, A. J., Cohen, D., Dhossche, D., Raffin, M., Wachtel, L., & Ferrafiat, V. (2023). Catatonia in neurodevelopmental disorders: Assessing catatonic deterioration from baseline. *Lancet Psychiatry, 10*, 228–234. https://doi.org/10.1016/s2215-0366(22)00436-9

Haynes, S. N., Richard, D. C., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*, 238–247.

Hedley, D., Uljarević, M., Foley, K. R., Richdale, A., & Trollor, J. (2018). Risk and protective factors underlying depression and suicidal ideation in autism spectrum disorder. *Depression and Anxiety, 35*, 648–657. https://doi.org/10.1002/da.22759

Hinze, E., Paynter, J., Dargue, N., & Adams, D. (2024). The Presentation of depression in depressed autistic individuals: A systematic review. *Review Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s40489-024-00480-z

Hirvikoski, T., Boman, M., Chen, Q., D'Onofrio, B. M., Mittendorfer-Rutz, E., Lichtenstein, P., Bölte, S., & Larsson, H. (2020). Individual risk and familial liability for suicide attempt and suicide in autism: A population-based study. *Psychological Medicine, 50*(9), 1463–1474. https://doi.org/10.1017/s0033291719001405

Hollo, A., Wehby, J. H., & Oliver, R. M. (2014). Unidentified language deficits in children with emotional and behavioral disorders: A meta-analysis. *Exceptional Children, 80*, 169–186. https://doi.org/10.1177/001440291408000203

Hossain, M. M., Khan, N., Sultana, A., Ma, P., McKyer, E. L. J., Ahmed, H. U., & Purohit, N. (2020). Prevalence of comorbid psychiatric disorders among people with autism spectrum disorder: An umbrella review of systematic reviews and meta-analyses. *Psychiatry Research, 287*, 112922. https://doi.org/10.1016/j.psychres.2020.112922x

Hudson, C. C., Hall, L., & Harkness, K. L. (2019). Prevalence of depressive disorders in individuals with autism spectrum disorder: A meta-analysis. *Journal of Abnormal Child Psychology, 47*, 165–175. https://doi.org/10.1007/s10802-018-0402-1

Jaffee, S. R., Moffitt, T. E., Caspi, A., Fombonne, E., Poulton, R., & Martin, J. (2002). Differences in early childhood risk factors for juvenile-onset and adult-onset depression. *Archives of General Psychiatry, 59*, 215–222.

Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, Williamson D, Ryan N. (1997). Schedulefor Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version(K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry*, *36*(7):980−988. https://doi.org/10.1097/00004583-199707000-00021. PMID: 9204677.

Kim, S. Y., & Lecavalier, L. (2021). Depression in young autistic people: A scoping review. *Research in Autism Spectrum Disorders, 88*, 101841. https://doi.org/10.1016/j.rasd.2021.101841

Kovacs M. The Children's Depression, Inventory (CDI). (1985). Psychopharmacol Bull, 21(4):995−958. PMID: 4089116

Kovacs, M. (1992). Manual of the Children's Depression Inventory. Toronto: Multi-HeathSystemsKovacs M. The Children's Depression Inventory (CDI) Manual. North Tonawanda, NY:Multi-Health Systems, Inc; 1992

Kroenke K, Spitzer RL, Williams JB. (2001). The PHQ-9: validity of a brief depression severitymeasure. J Gen Intern Med. 16(9):606−613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x. PMID: 11556941; PMCID: PMC1495268.

Lai, M. C., Kassee, C., Besney, R., Bonato, S., Hull, L., Mandy, W., Szatmari, P., & Ameis, S. H. (2019). Prevalence of co-occurring mental health diagnoses in the autism population: a systematic review and meta-analysis. *The Lancet Psychiatry, 6*(10), 819–829.

Leyfer, O. T., Folstein, S. E., Bacalman, S., Davis, N. O., Dinh, E., Morgan, J., Tager-Flusberg, H., & Lainhart, J. E. (2006). Comorbid psychiatric disorders in children with autism: Interview

development and rates of disorders. *Journal of Autism and Developmental Disorders, 36*(7), 849–861.

Magnuson, K. M., & Constantino, J. N. (2011). Characterization of depression in children with autism spectrum disorders. *Journal of Developmental and Behavioral Pediatrics, 32*, 332–340. https://doi.org/10.1097/DBP.0b013e318213f56c

Mandell, D. S. (2008). Psychiatric hospitalization among children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*, 1059–1065. https://doi.org/10.1007/s10803-007-0481-2

Mazefsky, C. A., Borue, X., Day, T. N., & Minshew, N. J. (2014). Emotion regulation patterns in adolescents with high-functioning autism spectrum disorder: Comparison to typically developing adolescents and association with psychiatric symptoms. *Autism Research, 7*, 344–354. https://doi.org/10.1002/aur.1366

Mazefsky, C. A., Kao, J., & Oswald, D. P. (2011). Preliminary evidence suggesting caution in the use of psychiatric self-report measures with adolescents with high-functioning autism spectrum disorders. *Research in Autism Spectrum Disorder, 5*, 164–174. https://doi.org/10.1016/j.rasd.2010.03.006

Mazurek, M. O., & Kanne, S. M. (2010). Friendship and internalizing symptoms among children and adolescents with ASD. *Journal of Autism and Developmental Disorders, 40*, 1512–1520. https://doi.org/10.1007/s10803-010-1014-y

Mazzone, L., Postorino, V., De Peppo, L., Fatta, L., Lucarelli, V., Reale, L., Giovagnoli, G., & Vicari, S. (2013). Mood symptoms in children and adolescents with autism spectrum disorders. *Research in Developmental Disabilities, 34*(11), 3699–3708.

Menezes, M., Harkins, C., Robinson, M. F., & Mazurek, M. O. (2020). Treatment of depression in individuals with autism spectrum disorder: A systematic review. *Research in Autism Spectrum Disorders, 78*, 101639. https://doi.org/10.1016/j.rasd.2020.101639

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & De Vet, H. C. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research, 19*(4), 539–549.

Newell, V., Phillips, L., Jones, C., Townsend, E., Richards, C., & Cassidy, S. (2023). A systematic review and meta-analysis of suicidality in autistic and possibly autistic people without co-occurring intellectual disability. *Molecular Autism, 14*, 12. https://doi.org/10.1186/s13229-023-00544-7

O'Connor, S. J., Hewitt, N., Kuc, J., & Orsini, L. S. (2023). Predictors and risk factors of treatment-resistant depression: A systematic review. *The Journal of Clinical Psychiatry, 85*(1), 50375. https://doi.org/10.4088/JCP.23r14885

Oliphant, R. Y. K., Smith, E. M., & Grahame, V. (2020). What is the prevalence of self-harming and suicidal behaviour in under 18s with ASD, with or without an intellectual disability? *Journal of Autism and Developmental Disorders, 50*, 3510–3524. https://doi.org/10.1007/s10803-020-04422-6

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. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ, 372*, n71.

Pandolfi, V., Magyar, C. I., & Norris, M. (2014). Validity study of the CBCL 6–18 for the assessment of emotional problems in youth with ASD. *Journal of Mental Health Research Intellectual Disabilities, 7*, 306–322. https://doi.org/10.1080/19315864.2014.930547

Pezzimenti, F., Han, G. T., Vasa, R. A., & Gotham, K. (2019). Depression in youth with autism spectrum disorder. *Child and Adolescent Psychiatric Clinics of North America, 28*, 397–409. https://doi.org/10.1016/j.chc.2019.02.009

Pilunthanakul, T., Goh, T. J., Fung, D. S. S., Sultana, R., Allen, J. C., & Sung, M. (2021). Validity of the patient health questionnaire 9-item in autistic youths: A pilot study. *BMC Psychiatry, 21*, 564. https://doi.org/10.1186/s12888-021-03556-w

Popow, C., Ohmann, S., & Plener, P. (2021). Practitioner's review: Medication for children and adolescents with autism spectrum disorder (ASD) and comorbid conditions. *Neuropsychiatrie, 35*, 113–134. https://doi.org/10.1007/s40211-021-00395-9

Rai, D., Heuvelman, H., Dalman, C., Culpin, I., Lundberg, M., Carpenter, P., & Magnusson, C. (2018). Association between autism spectrum disorders with or without intellectual disability and depression in young adulthood. *JAMA Network Open, 1*(4), e181465–e181465. https://doi.org/10.1001/jamanetworkopen.2018.1465

Rice, F., Riglin, L., Lomax, T., Souter, E., Potter, R., Smith, D. J., Thapar, A. K., & Thapar, A. (2019). Adolescent and adult differences in major depression symptom profiles. *Journal of Affective Disorders, 243*, 175–181. https://doi.org/10.1016/j.jad.2018.09.015

Santomauro, D., Sheffield, J., & Sofronoff, K. (2016). Depression in adolescents with ASD: A pilot RCT of a group intervention. *Journal of Autism and Developmental Disorders, 46*, 572–588. https://doi.org/10.1007/s10803-015-2605-4

Sterling, L., Renno, P., Storch, E. A., Ehrenreich-May, J., Lewin, A. B., Arnold, E., Lin, E., & Wood, J. (2015). Validity of the Revised Children's Anxiety and Depression Scale for youth with autism spectrum disorders. *Autism, 19*(1), 113–117.

Terwee, C. B., Jansma, E. P., Riphagen, I. I., & de Vet, H. C. W. (2009). Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research, 18*, 1115–1123.

Thapar, A., Collishaw, S., Pine, D. S., & Thapar, A. K. (2012). Depression in adolescence. *The Lancet, 379*, 1056–1067. https://doi.org/10.1016/S0140-6736(11)60871-4

Thapar, A., Livingston, L. A., Eyre, O., & Riglin, L. (2023). Practitioner review: Attention-deficit hyperactivity disorder and autism spectrum disorder—the importance of depression. *Journal of Child Psychology and Psychiatry, 64*, 4–15. https://doi.org/10.1111/jcpp.13678

Thomas, H. R., Sirsikar, A., & Eigsti, I. M. (2024). Brief report: Convergence and discrepancy between self- and informant-reported depressive symptoms in young autistic adults. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-023-06230-0

Thurm, A., Halladay, A., Mandell, D., Maye, M., Ethridge, S., & Farmer, C. (2022). Making research possible: barriers and solutions for those with ASD and ID. *Journal of Autism and Developmental Disorders, 52*, 4646–4650. https://doi.org/10.1007/s10803-021-05320-1

Uljarević, M., Richdale, A. L., McConachie, H., Hedley, D., Cai, R. Y., Merrick, H., Parr, J. R., & Le Couteur, A. (2018). The hospital anxiety and depression scale: Factor structure and psychometric properties in older adolescents and young adults with autism spectrum disorder. *Autism Research, 11*(2), 258–269.

Volkmar, F., Siegel, M., Woodbury-Smith, M., King, B., McCracken, J., & State, M. (2014). Practice parameter for the assessment and treatment of children and adolescents with autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 53*, 237–257. https://doi.org/10.1016/j.jaac.2013.10.013

Westby, C. (2022). Nature and effects of autobiographical memory issues in persons with autism spectrum disorders. *Neuropsychiatric Disease and Treatment, 18*, 2279–2293. https://doi.org/10.2147/ndt.S332521

Zigmond AS, Snaith RP. (1983). The hospital anxiety and depression scale. Acta Psychiatr Scand. 67(6):361−370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x. PMID: 6880820.