ANNA EAST BA (Hons) MSc

Establishing global eligibility criteria for a diagnosis of autism in a sporting context: availability of assessments and views of an expert panel

Section A: What measures are available for the assessment of autism? A systematic meta-review of screening and diagnostic assessments for use with older children, adolescents, and adults.

ACCURATE WORD COUNT: 7983 (127)

Section B: Increasing access to competitive sport: A Delphi study exploring the views of an expert panel on a method for establishing global eligibility criteria for a diagnosis of autism in a sporting context.

ACCURATE WORD COUNT: 7976 (481)

OVERALL WORD COUNT: 15,959 (608)

Submitted in partial fulfilment of the requirements of Canterbury Christ Church University for the degree of Doctor of Clinical Psychology

MAY 2023

SALOMONS INSTITUTE FOR APPLIED PSYCHOLOGY

CANTERBURY CHRIST CHURCH UNIVERSITY

Acknowledgments

Thank you so much to my research supervisor, Jan Burns for your guidance and support over throughout the project. Your passion and expertise in this area has been invaluable. Thank you for being so responsive whenever I had questions. Thank you to my external supervisor Mark Murphy for your thoughtful comments.

My sincere thanks to Matt Lynch for reading through each of the surveys. Your helpful comments and thoughts on the process were very appreciated, and I enjoyed sharing the project with you. Thank you to each one of the participants who took part in this study, your contributions were so valuable, and I was very grateful for the time you spent responding to the surveys.

Thank you to my partner, Robin, and my family for always supporting and encouraging me.

Summary of the Major Research Project

Section A

A systematic meta-review was conducted to examine review papers on autism assessment measures for older children, adolescents, and adults. Twenty review papers were summarised and evaluated using a narrative synthesis. Twenty-eight screening and fifteen diagnostic measures for autism were identified from the reviews, and the characteristics and psychometric properties of these measures were critically evaluated. There is a need for a more thorough investigation of a wider range psychometric properties for the screening and diagnostic measures. Overall, there was little information within the reviews about where measures have been validated or what languages they have been translated into which presents an opportunity for further research.

Section B

This study aimed to examine the views of an expert panel on a method for establishing global eligibility criteria for a diagnosis of autism in a sporting context. Twenty-seven international participants took part in a three-round Delphi panel using online surveys. The results of the study showed that there was high consensus around a gold standard process of eligibility. There were lower levels of consensus and agreement around whether there should be an alternative process for countries that are unable to access the gold standard. Key challenges and barriers were identified including social and cultural differences, attention to co-morbidity and the heterogeneity of autism. The need for further research to explore how autism impacts performance during sports competition was discussed.

Contents

Section A List of Tables and Figures

List of Tables

List of Figures

Section B List of Tables and Figures

List of Tables

List of Figures

Section C: Appendices of Supporting Material Contents

ANNA EAST BA (Hons) MSc

SECTION A:

What measures are available for the assessment of autism? A systematic meta-review of screening and diagnostic assessments for use with older children, adolescents, and adults.

Word Count: 7983 (127)

A thesis submitted in partial fulfilment of the requirements of Canterbury Christ Church University for the degree of Doctor of Clinical Psychology

APRIL 2023

SALOMONS INSTITUTE FOR APPLIED PSYCHOLOGY

CANTERBURY CHRIST CHURCH UNIVERSITY

Abstract

Considering the array of review papers that focus on evaluating autism assessment measures for older children, adolescents and adults, there is a need to summarise the existing evidence to highlight consistencies and differences in the research. A meta-review was conducted to evaluate the quality of these reviews using a narrative synthesis. 20 review papers were included, compromising 10 systematic and 10 non-systematic reviews. The quality of the review papers was generally found to be high, but there were concerns around critical appraisal, data extraction and addressing publication bias. From these reviews 28 screening and 15 diagnostic measures were identified for autism spectrum disorder (ASD), and the characteristics and psychometric properties of these measures were synthesised and critically evaluated. The ADOS-2 and ADI-R were the measures most recommended by the reviews, however there was variability in reports of validity and reliability. There is a need for a more thorough investigation of a wider range psychometric properties for ASD screening and assessment measures. Overall, there was little information within the reviews about where ASD measures have been validated or what languages they have been translated into. This presents an important direction for future research.

Keywords: Autism; Assessment; Meta-review; Screening; Diagnostic

Introduction

Autism spectrum disorder (ASD), or autism spectrum condition (ASC), is a neurodevelopmental condition defined by deficits in social communication and interaction, the presence of repetitive and stereotyped behaviour, and sensory sensitivities (American Psychiatric Association [APA], 2013). ASD is used throughout this meta-review to ensure consistency, as this is the most common term used in the language of autism assessment measures. With increasing knowledge about autism, standardized diagnostic criteria for ASD have been published and refined over the years (Rosen et al., 2021). The most updated and widely used criteria for ASD are detailed in the Diagnostic and Statistical Manual of Mental Disorders Fifth Edition (DSM-V; APA 2013) and the International Classification of Diseases 11th Edition (ICD-11; WHO 2018). Standardised diagnostic criteria create a foundation for assessment measures that help to identify ASD (Rosen et al., 2021).

Prevalence of Autism

Research has shown an increasing trend in the prevalence of ASD (Salari et al., 2022). This has been contributed to increased awareness, changes in diagnostic criteria, and shifts in research methods, alongside the potential of a true increase in prevalence (Durkin & Wolfe, 2020). However, there is a wide variation in prevalence rates globally (Salari et al., 2022). A recent systematic review update by Zieden (2022) found that autism prevalence ranges from 1 in 10,000 to 436 in 10,000 across the world. Estimates in prevalence have been consistently higher in Western Countries, such as the US (Christensen et al., 2018). This variety likely reflects the complex and dynamic interactions between patterns of community awareness, service capacity, help seeking, sociodemographic factors, as well differences in access to valid and reliable assessments (Rosen et al., 2021). While standardized measures have been used across different countries (Marlow et al., 2019) and diverse populations (Harrison et al., 2017), review studies have shown there are vast differences in access to culturally adapted,

translated, valid and reliable assessments (Soto et al., 2015). Furthermore, less research originates from low- and middle-income countries (LMICs), resulting in their underrepresentation within ASD literature (Franz et al., 2017).

Assessment of Autism

Assessment of autism is important to facilitate access to support and services (Gabbay-Dizdar, 2022). ASD can have varying impacts on an individual's life, ranging from mild to profound impairments that can be associated with difficulties in adaptive functioning and everyday activities (Elder et al., 2017). Individuals with autism are more likely to have co-occurring physical health conditions (Curtin et al., 2014), intellectual disabilities, language impairments, and mental health conditions (Lai et al., 2019; Matson et al., 2013). Therefore, access to, and use of, valid and reliable assessment measures is important to accurately identify ASD and to enable access to support (Elder, 2017).

ASD is a heterogeneous condition with many differences in the presentation of symptoms, which makes the diagnostic process particularly complex (Rosen et al., 2021). In the absence of biological methods (Arnett et al., 2019), the identification of ASD remains a multidisciplinary and clinical assessment, requiring information from multiple sources (McCarty & Frye, 2020). Diagnosis of ASD utilises psychometric assessments and developmental histories, which require extensive clinical expertise and training as well as a focus on differential diagnoses (Kamp-Becker et al., 2021).

There are many different assessment measures that exist for the identification of ASD (Charman & Gotham, 2013). Current measures rely on self-, or informant-report questionnaires, observation-based measures, and standardised structured interviews (Carpenter, 2012). Diagnostic measures are more comprehensive, time-consuming, and are delivered by trained professionals as part of a multi-disciplinary assessment (NICE, 2012).

Measures designed for screening are used to justify a more in-depth assessment, or as a first step in the diagnostic process (Baer & Blais, 2010). Some countries, such as the UK, have produced national assessment guidelines (e.g. NICE, 2012), and the current diagnostic gold standard is considered to be a multi-disciplinary assessment using the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview Revised (ADI-R) (Kamp-Becker et al., 2021). However, even gold-standard diagnostic tools show a large variation in diagnostic capabilities (Lefort-Besnard et al., 2020).

Evaluating Assessment Measures

ASD assessment measures have to navigate the complexity and heterogeneity of symptoms in the presentation of ASD, alongside the co-occurrence of intellectual disabilities and other medical conditions described above (Lai et al., 2019). Assessments are also influenced by differences in nations and cultures, as views regarding appropriate behaviours and normal child development, differ across the world (Matson et al., 2017). The creation and evaluation of ASD assessment measures has predominantly been conducted in the West, within English-speaking industrialised countries (Wallis & Pinto-Martin, 2008). More recent studies have looked at cultural adaption and translation of ASD measures in non-English speaking countries and regions (Soto et al., 2015; Al Maskari et al., 2018). However, there is a need for further research on where ASD measures have been translated and validated (Matson et al., 2017).

This heterogeneity has contributed to the development of a wide number of screening and diagnostic measures that can be used in different populations, contexts, and cultures (Soto et al., 2015). In addition to the heterogeneity of the diagnosis, recent research on gender difference is having a significant impact on how and when autism is recognised in females (Hull et al., 2020). While autism is more commonly diagnosed in males, research has suggested that females can express autism in ways which do not meet current diagnostic

criteria and are therefore not fully capture by assessment measures (Hull et al., 2020). This has placed criticism on existing tools and raised the need to review and revise existing measures (Loomes et al. 2017).

Diagnostic and screening measures for ASD have been evaluated according to their general characteristics, psychometric measurement properties, and generalizability to different populations (Falkmer et al., 2013). Within the ASD literature, there are both individual validation studies (e.g. Charman et al., 2007), and review studies that summarise and evaluate a number of different measures (e.g. Baghdadli et al., 2018). Due to the breadth of individual validation studies for ASD measures, systematic reviews synthesising information are particularly useful in guiding clinicians and researchers in the selection of the most appropriate ASD screening and diagnostic tools. There is an array of review studies that have been conducted covering a range of characteristics including type of assessment, age, culture, context, population, and co-morbidity with other conditions (e.g., Wigham et al., 2019; Levy et al., 2020; Wang et al., 2020). This includes different types of reviews, from systematic reviews with strict criteria (e.g., Hirota et al., 2018) to more narrative reviews (e.g., Charak & Stella, 2002), across a wide range of dates.

The Need for a Meta-Review

Considering the array of review papers that focus on evaluating ASD assessment measures, there is a need to explore how many of these reviews have been conducted, where they have been conducted, and what ASD measures they evaluate. This would help to summarise the existing evidence in this area, and bring together information about screening and diagnostic measures designed for different ages and different regions (Hennessy et al., 2019). Given the heterogeneity in the assessment of ASD, this would help to identify what assessments are available for screening and diagnosing ASD and what the quality of these assessments are. Summarising the findings of different reviews would highlight consistencies

and differences in the research about ASD assessment measures. This would further help to identify where there are gaps in knowledge and indicate the need for further research.

There is also a need to evaluate the quality of existing review papers and the evidence base for ASD assessment measures. Given the number of measures and individual studies, review papers are particularly useful at summarising the available evidence. However, it is important to consider the quality of these review papers, to make appropriate judgements regarding their findings and recommendations. Examining the quality of the review papers would further help to identify the strength of the overall evidence base and make recommendations for future research in this area. Evaluating the current evidence base for ASD assessment measures increases potential opportunities for cross-cultural research and interventions. This is important to help strengthen capacities for the assessment of ASD, which was identified as a global aim following a World Health Organisation (WHO) consultation in 2013 (WHO, 2013).

Conducting a meta-review presents the opportunity to identify and summarise a number of reviews in one place, and to evaluate the quality of these reviews. This is particularly useful where there are multiple reviews on one topic (Hennessy et al., 2019). This meta-review focussed on ASD measures for use with older children, adolescents, and adults above the age of six and without co-occurring intellectual disabilities. ASD measures for young children have a much greater focus on observational assessment, compared to the structured interviews and questionnaires used with older children and adults. With increased interest in early detection and intervention, there is already extensive research on evaluating ASD measures for infants and young children (Marlow et al., 2019). In comparison, less research has been conducted on ASD screening and diagnostic tools with older children, adolescents, and adults (Bagdadlhi et al., 2017; Hirota et al., 2018). Age six is generally considered to be the start of middle childhood and the end of ‘pre-school’ years (Kandice

Mah & Lee Ford-Jones, 2012). Limiting the age of assessment allowed this meta-review to go into enough depth to provide a comprehensive overview of ASD screening and diagnostic measures for use with older children, adolescents, and adults.

Summary of Aims

The aims of this meta-review were to:

1. Identify the reviews that have been published that evaluate autism screening and diagnostic measures for older children, adolescents, and adults.
2. Examine the quality of the included reviews.
3. Describe and examine the characteristics and psychometric properties of the identified screening and diagnostic measures.
4. Summarise and compare the conclusions of reviews about autism screening and diagnostic measures.

Method

The steps detailed in Hennessy et al. (2019)'s best practice guidelines for systematic meta-reviews were followed as far as was pragmatically possible. The aims were discussed in supervision to help define focus and scope of the meta-review (Hennessy et al., 2019). The term 'meta-review' was used instead of 'overview' or 'umbrella review', which could have implied broader scope or the synthesis of both primary studies and review literature.

Search Strategy

In February 2022, a literature search was conducted using the databases Medline, PsychInfo, Cochrane, British Education Index (EBSCO), Education Resources Information Centre (ERIC), and Australian Education Index. A wide range of databases were searched to ensure that relevant reviews were located, and to reduce the potential of selection bias

(Hennessy et al., 2019). The Cochrane database was included in the search as it a large database of reviews, and the British, American and Australian Education Indexes were searched as autism measures have been used in education-based settings (Thabtah & Peebles, 2019). An unrestricted date range was used to ensure that the meta-review could provide a comprehensive overview of the literature (Hennessy et al., 2019), and could consider the quality of reviews published at any time. This was important to capture information on older as well as more recent autism measures, and to gather information on how autism measures have been updated over time. Search terms and strategy are detailed in Tables 1 and 2.

Table 1

Search terms

Search Terms
Autis* OR Asperger* OR ASD OR ASC
AND
Assess* OR tool* OR measure* OR questionnaire* OR scale* OR test* OR screen* OR diagnos*
AND
Review* OR met-analys* OR summar*

Table 2

Search strategy

Search Strategy in Databases
In titles for Medline and PsychInfo
In titles, abstracts, and keywords for Cochrane
In titles and abstracts for the British Education Index (EBSCO) and the Australian Education Index
In titles for Education Resources Information Centre (ERIC)

The search was limited to human participants and articles published in English language. No date limitations were used for the search, to ensure that the meta-review would provide a comprehensive overview of the research literature. Practice searches were conducted to craft the search terms and strategy, alongside discussion with the research supervisor (Hennessy et al., 2019). Manual searching of the included papers' reference lists was conducted, alongside using the search terms in Google Scholar to find any review papers that might have been missed. A PRISMA diagram was produced showing the screening and selection processes (Figure 1; Moher et al., 2009).

Inclusion and Exclusion Criteria

Papers were screened against the inclusion and exclusion criteria detailed in Table 3.

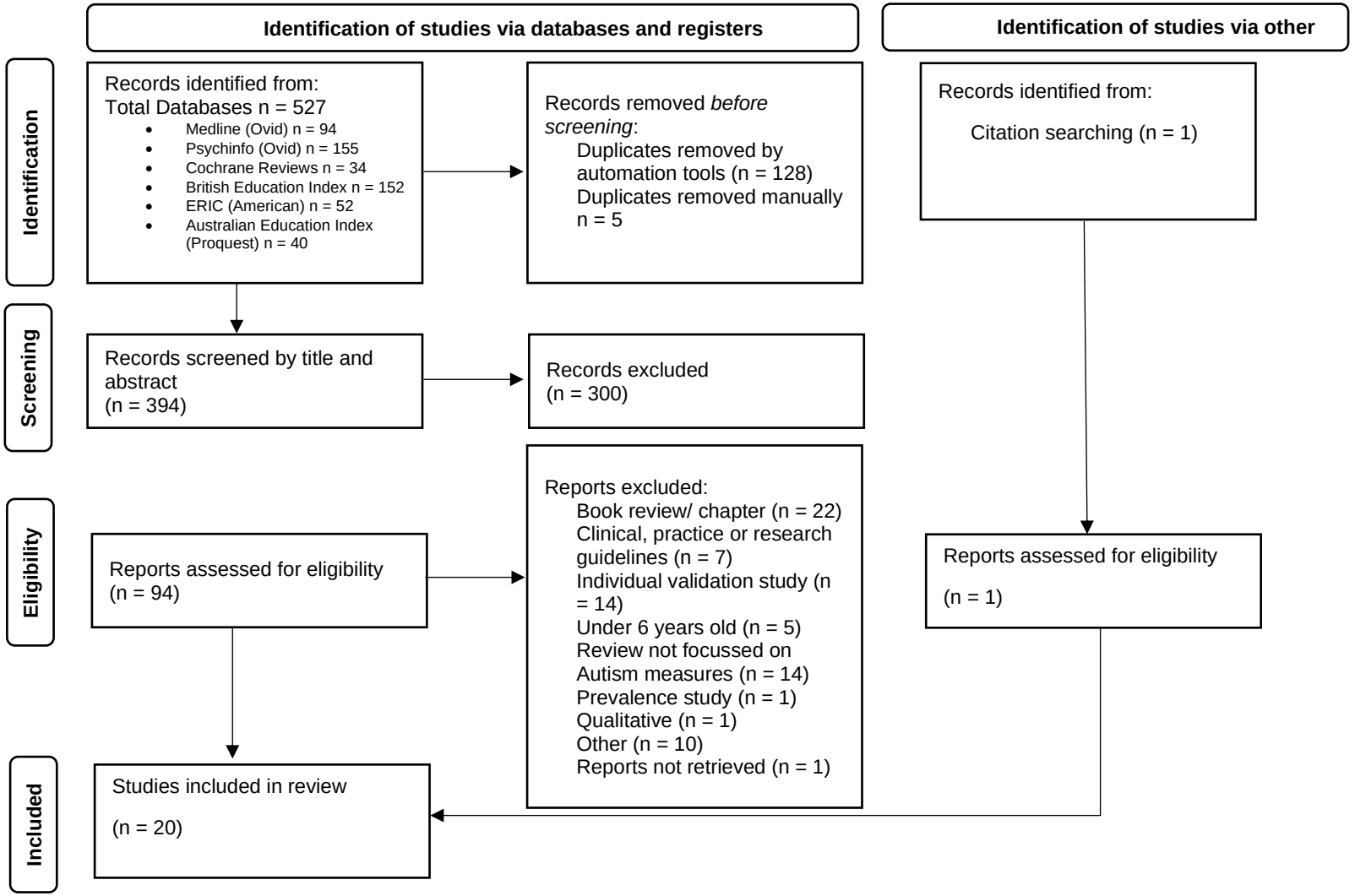
Table 3

Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
• Journal published review, summary or meta-analysis study that has a predominant focus of reviewing multiple measures or assessment tools for identifying or diagnosing autism	• Individual validation and research studies for autism assessment measures *or* reviews that focus on one or two individual autism assessment measures (such as comparative reviews)
• Autism measures or assessment tools designed for older children, adolescents, and adults	• Book reviews or book chapters
• Autism measures or assessment tools designed for screening or diagnostic use	• Qualitative research reviews. This was defined as either: o Descriptive reviews or summaries which provided no evaluative data o Reviews that only included primary studies with qualitative data
• Autism measures or assessment tools designed for any point on the autism spectrum (e.g., including measures designed for Asperger's)	• Review papers not published in English
• Reviews focused on any region or country in the world	• Reviews that had a primary focus on: o Autism assessment measures designed for just for use with co-occurring intellectual disabilities and not designed for use with mean normal intelligence o Autism assessment measures for infants or young children aged 6 and under
	• Dissertations, theses, or reviews that were not peer-reviewed
	• Reviews that just included measures for any type of atypical development. Instead, only reviews that included assessment measures specifically designed to screen for or identify autism were included.

Figure 1

PRISMA diagram detailing the screening and selection processes



Data Extraction

The titles and abstracts were reviewed against the inclusion and exclusion criteria. Titles and abstracts that met the inclusion criteria or where there was uncertainty, were set aside for full text review. The full texts of these papers were then reviewed against the inclusion criteria. The final included reviews were separated into systematic reviews and non-systematic reviews. Reviews were categorised as systematic if the review paper described the type of review as systematic, or if the paper stated that it followed PRISMA guidelines (Moher et al., 2009; Page et al., 2021).

A data extraction tool was developed based on the purpose and aims of the meta-review (Appendix A; Hennessy et al., 2019). This was piloted on several of the included papers, and then discussed with the research supervisor. Data was extracted using the tool from the included reviews and was organised into the detailed sections in Table 4.

Table 4

Areas for data extraction

Data Extracted from the Included Reviews:
Information about each systematic review (author(s); year; type of measures included, population, country/ region, short summary, main findings);
Information about each non-systematic review (author(s); year; type of measures included, population, country/ region, short summary; main findings);
Screening measures (author(s); year; age group, administration time, administration mode e.g., questionnaire, short summary);
Diagnostic measures (author(s); year; age group, administration time, administration mode e.g., questionnaire, short summary);
Findings related to the psychometric properties of screening and diagnostic measures (internal consistency; reliability; validity; sensitivity; specificity), and information about the quality of the primary study evidence base.

Calculation of Overlap

The overall Corrected Covered Area (CCA) was calculated using the steps laid out in Hennessy and Johnson (2020), by generating a citation matrix (Appendix B). Determining the overall degree of overlap in a meta-review is not enough, as it may conceal overlap between sub-clusters of reviews (Hennessy & Johnson, 2020). Therefore, the citation matrix was examined, and the CCA was then calculated between any two reviews that had included more than two of the same primary studies. High-quality reporting is a pre-requisite for calculation of overlap and therefore, analysis of overlap was not conducted for any reviews that did not have a complete list of included primary studies (Pieper et al., 2014).

Data Analysis

In a meta-review, it is important to consider the methodological limitations of the included reviews as well as the quality of the primary studies that they contain, to determine the overall quality of the existing evidence (Hennessy et al., 2019). The quality of the included reviews was assessed using two different standardised quality checklists. The quality of systematic reviews was assessed using the Joanna Briggs Institute (JBI) Checklist for Systematic Reviews and Research Synthesis (Appendix C; Aromataris et al., 2015), which critically appraises methodological quality and risk of bias. The JBI checklist compromises 11 questions, scoring ‘Yes’, ‘No’, ‘Unclear’, or ‘Not Applicable’ for each question. The total number of questions rated ‘Yes’ was presented. As all the non-systematic review papers used narrative methods to synthesise data, the quality of non-systematic reviews was assessed using the Scale for the Assessment of Narrative Review Articles (SANRA) scale (Appendix D; Baethge et al., 2019). This is currently the only tool specifically designed to assess the quality of narrative reviews (Baethge et al., 2019). The SANRA scale is comprised of six questions that reflect six aspects of quality, rated on a scale scored zero (not addressed at all), one (partially addressed) or two (fully addressed). The

scores of all six questions were summed to give an overall score. As the aim of this meta-review was to provide an overview of the literature, none of the included reviews were excluded based on quality scores (Hennessy et al., 2019). Quality scores were used to inform critical thinking about the quality of the literature and are considered further in the results and discussion sections. To ensure validity of the quality ratings, 30% of the included reviews were critically appraised by a second reviewer, and discrepancies were resolved by conversation with a third reviewer.

The characteristics and psychometric properties of screening and diagnostic measures were identified from the included reviews. These were summarised following the cyclical process steps for narrative synthesis in meta-reviews from Hennessy et al., (2019; Table 5).

Table 5

Narrative Synthesis

Narrative Synthesis Process	
1. Preparation phase:	The reviews were read in depth multiple times to facilitate data immersion and make sense of the data (Elo & Kyngas, 2008).
2. Organisation phase:	Notes and headings were written for each of the review articles to describe as many aspects of the reviews as possible and help create categories (Appendix E).
3. Abstraction phase:	A general description of the data was formulated by grouping categories and features together. Appropriate presentation of these categories was decided.

Harvest plots (Ogilvie et al., 2008) were produced to visually display psychometric properties for the screening and diagnostic measures that had been identified from the included reviews. Harvest plots are a method for graphically displaying evidence from complex and diverse studies, and are useful where data has been reported in different formats (Crick et al., 2015; Crowther et al., 2011). Each review paper was represented by a bar

positioned according to a category. The bars could be ‘visually weighted’ (by height or width) and annotated to highlight study and outcome characteristics, as well as study quality (Crowther et al., 2011; Ogilvie et al., 2008). Harvest plots have been used in meta-reviews both with and without meta-analysis (Crick et al., 2015), and in systematic reviews to synthesise and display the results of psychometric properties (Edwards et al., 2020). Psychometric property data was first extracted from the review papers and synthesised into tables (Appendices F-G). Harvest plots were then produced and combined into a matrix, following the steps in Ogilvie et al. (2008).

Results

Reference Selection

From the literature search, 394 references were identified after duplicates were removed. 300 of these references were excluded at title and abstract review and 94 full text articles were assessed. Of these, 75 did not meet the inclusion criteria with the main exclusion factors being book chapters or book reviews, and individual validation studies. Reviews of autism that only included a small section on autism measures, such as reviews summarising diagnostic criteria or theories of autism, were also excluded as the focus of these reviews was not on autism measures. Two reviews had inclusion criteria of age seven and under, but these were excluded as the majority of assessments included were designed for use infants and young children below six years (Al Maskari et al., 2018; Marlow et al., 2019). One further reference was identified through citation searching which met the inclusion criteria. Therefore, 20 review papers were included in this meta-review.

Summary of the Included Review Papers

Out of the 20 included review studies, 10 were categorised as systematic reviews and 10 were categorised as non-systematic reviews.

Systematic Reviews

Table 6 presents the general characteristics of the 10 systematic reviews that were identified. Overall, there were more systematic reviews looking at screening measures than diagnostic measures for autism. Five of the reviews focussed solely on screening measures for autism (Hirota et al., 2018; Levy et al., 2020; Soto et al., 2015; Stewart & Lee 2017; Wang et al., 2020), with Levy et al., (2020) just including screening measures that can be used in primary care or primary care-like settings. Four of the systematic reviews included both screening and diagnostic measures (Backes et al., 2014; Baghdadli et al., 2017; Sun et al., 2013; Wigham et al., 2019), and just one of the reviews focussed solely on diagnostic measures but this review did include a screening measure that had the potential to be diagnostic (Falkmer et al., 2013). The systematic reviews included a range of ages. Two of the reviews just looked at children (Levy et al., 2020; Sun et al., 2013), two of the reviews looked at children and adolescents (Soto et al., 2015; Wang et al., 2020) and two of the reviews only included measures designed for use with adults (Baghdadli et al., 2017; Wigham et al., 2019). The other four reviews included autism assessments for use with children, adolescents, or adults (Backes et al., 2014; Falkmer et al., 2013; Hirota et al., 2018; Stewart & Lee 2017).

Some of the systematic reviews only included autism assessments that were designed for use in a certain country or type of region. This included Brazil (Backes et al., 2014), Mainland China and the surround regions (Sun et al., 2013; Wang et al., 2020), and Low- and Middle-income countries as defined by the World Bank, such as Mexico and Iran (Stewart & Lee 2017). The other systematic reviews did not specify a particular country or region, but these were limited to studies published in English, the Western World, or countries highly rated on the human development index. Finally, one systematic review did not specify a region but focussed on cultural adaptations of existing autism measures (Soto et al., 2015).

Table 6

Summary of systematic review papers

Authors	Type of measures	Population	Region or Country	Short Summary	Summary of Main Findings
Backes et al., 2014	Screening and diagnostic	Children, adolescents, and adults	Brazil	Reviews assessment measures for ASD in the Brazilian population and assesses the psychometric properties of these measures. Examines the quality of the studies conducted in Brazil that investigate the psychometric properties and comments of the suitability and appropriateness of these measures.	Identified six ASD assessment instruments studied in the Brazilian population. All instruments that assessed internal consistency showed adequate values. The Autism Diagnostic Interview-Revised (ADI-R) and the Childhood Autism Rating Scale (CARS) adaptations showed satisfactory inter-rater reliability and test-retest indices. There are still no specific ASD diagnostic tools available for use in Brazil. There was lack of information on copyrights for some of the instruments.
Baghdadli et al., 2017	Screening and diagnostic	Adults	Any but only articles published in English and French	Reviews assessment measures for screening and diagnosis of ASD in adults without intellectual disabilities. Examines the psychometric measurement properties of these measures and assesses diagnostic accuracy.	Identified nine measures (and short versions). There was unsatisfactory evidence within the included studies due to poor methodological quality and small sample sizes. None of the measures had satisfactory psychometric properties supported by strong evidence. The Autism Spectrum Quotient (AQ-50), Autism Spectrum Quotient-Short Form (AQ-S), Ritvo Asperger and Autism Diagnostic Scale-Revised (RAADS-R), and Ritvo Asperger and Autism Diagnostic Scale-Short Form (RAADS-14) had the most satisfactory values for psychometric properties.
Falkmer et al., 2013	Diagnostic (but includes screening measures with the potential to be diagnostic)	Children, adolescents, and adults	Western world and published in English	Examines the psychometric properties of ASD diagnostic measures for all ages to identify the optimal diagnostic instrument.	Assessed 17 measures for ASD. There was a limited evidence base for most of the measures due to a lack of high-quality studies. The Autism Diagnostic Interview-Revised (ADI-R) and Autism Diagnostic Observation Schedule (ADOS) scored highest on sensitivity and specificity with the strongest evidence base. The highest levels of accuracy were achieved with using the ADI-R and ADOS in combination. More research is needed on the use of these measures together.
Hirota et al., 2018	Screening	Children above the age of 4,	Any but only included articles	Reviews the validity and psychometric properties of screening measures for ASD in older children and adults. Provides recommendations about the use of ASD measures in a variety of settings.	Identified 11 screening tools. Only the Autism-Spectrum Quotient (ASQ), the Social Communication Questionnaire (SCQ), and the Social Responsiveness Scale (SRS) were examined by multiple studies. These

		adolescents and adults	published in English		three measures may assist in differentiating ASD from other neurodevelopmental conditions. There is a lack of research with young adult populations and a lack of data about ASD measures that can be used globally in areas with fewer resources and knowledge.
Levy et al., 2020	Screening	Children up to 12 in primary care (PC) or primary care-like settings	Studies in any country rated as "very high human development" on the United Nations Development Program's International Human Development Index	Reviews screening measures for ASD for children ages up to 12 used in primary care and primary care like settings. Examines the psychometric properties of these screening measures.	Identified seven screening measures but there was a lack of research evaluating these screening measures in primary care for children in the 6 to 12 age range. Limited evidence evaluating the sensitivity, specificity, and negative predictive value of instruments was available.
Soto et al., 2015	Screening	Children and adolescents up to 18	Any but focussed on cultural adaptations of existing tools (articles could be published in languages other than English)	Reviews ASD screening measures that have been culturally adapted to different cultures and countries. Assesses the adaptation process against cultural adaptation guidelines. Examines the psychometric properties of the adapted ASD screening measures.	Identified nine adapted measures. The cultural adaptation process was not described in enough detail and did not follow recommended guidelines for most measures. There was variety in the level of psychometrics reported. Only nine studies reported internal consistency, with less reporting test–retest reliability and factor structure. There was limited evidence for convergent or concurrent validity of adapted measures. Differences in the psychometric properties of the original and adapted measures were found by most studies. There is a need for further normative data to increase the utility of translated or adapted measures.
Stewart & Lee 2017	Screening	Children, adolescents, and adults	Low- and middle-income countries (LMICs) as defined by the World Bank	Reviews ASD screening measures that are used in LMICs. Assesses the study design and screening methodology for ASD measures used in LMICs.	Assessed 18 screening measures used in LMICs. There was significant variation in study design, screening methodology, and population characteristics in the included studies. There was a wide range of cut-off points, sensitivities, and specificities reported within screening instruments. Few studies reported formal adaptation procedures.

Sun et al., 2013	Screening and diagnostic	Children	Mainland China	Reviews the ASD criteria and measures that have been used for ASD diagnosis in mainland China. Examines the psychometric properties of these measures, including validity and reliability.	Eight screening instruments and two diagnostic instruments were identified. The most frequently used instruments in mainland China are the Clancy Autism Behaviour Scale (CABS), the Autism Behaviour Checklist (ABC) and the Childhood Autism Rating Scale (CARS). The Autism Diagnostic Interview (ADI) has not been thoroughly adopted in mainland China. There is a lack of consistency in the methodology within the included studies. There was large variation used in the cut-offs for each measure. There needs to be further validation of standardised instruments in mainland China.
Wang et al., 2020	Screening	Children and adolescents under 18 years old	Mainland China and surrounding regions	Reviews ASD screening measures currently used in mainland China and surrounding regions. Examines the strengths and limitations of these ASD screening measures. Provides recommendations for ASD screening in Chinese-speaking countries.	Identified 10 screening measures. Screening instruments showed fair to good sensitivity and specificity. Reported psychometric properties were encouraging but were not based on sufficient data which limited the findings. Studies varied greatly in the extent of psychometric analyses and reported autism spectrum disorder prevalence.
Wigham et al., 2019	Screening and diagnostic	Adults	Any but only included articles published in English	Reviews ASD diagnostic measures that have been published since the National Institute for Health and Care Excellence (NICE) guidelines update. Assesses the quality of the ASD measures and included studies. Provides recommendations about the use of ASD measures in adults.	Assessed 12 screening and diagnostic measures. There is limited evidence for the use of structured questionnaires for assessing ASD. The Autism Diagnostic Interview-Revised (ADI-R) and Autism Diagnostic Observation Schedule (ADOS) had mixed results for sensitivity and specificity. The Autism Mental State Examination (AMSE) showed promising evidence as a diagnostic tool. Further research is needed to investigate the accuracy of the 12 measures in this review.

NR: Not Reported

Non-Systematic Reviews

Table 7 presents the general characteristics of the 10 non-systematic reviews that were identified. Five of the non-systematic reviews only included screening measures (Campbell 2005; Norris & Lecavalier 2010; Sappok et al., 2015; Stoesz et al., 2011; Thabtah & Peebles 2019), four looked at both screening and diagnostic measures (Charak & Stella 2002; Klose et al., 2012; Matson et al., 2007; Parks 1983), and one non-systematic review solely focussed on diagnostic measures (Morgan 1988). Only one of the non-systematic reviews looked at autism measures just for adults aged 18 and above (Stoesz et al., 2011), with five of the reviews assessing measures for children and adolescents (Campbell 2005; Charak & Stella 2002; Morgan 1988; Norris & Lecavalier 2010; Parks 1983). The remaining four reviews included measures for all ages (Klose et al., 2012; Matson et al., 2007; Sappok et al., 2015; Thabtah & Peebles 2019). None of the non-systematic reviews included information specifying which region or country the measures were designed for use in, however the review by Sappok et al. (2015) specified the inclusion of only English and German screening measures.

Quality of the Included Review Papers

Table 8 presents the critical appraisal of the 10 systematic review papers according to the JBI Checklist (Aromataris et al., 2015). Most of the systematic reviews were rated as ‘Yes’ to seven or more of the 11 areas included in the checklist. Three of the systematic reviews scored higher with nine or more areas rated as ‘Yes’ on the JBI checklist (Hirota et al., 2018; Levy et al., 2020; Wigham et al., 2019), and one of the systematic reviews scored lower with just five areas rated as ‘Yes’ (Falkmer et al., 2013). There were concerns highlighted for many of the systematic reviews around whether appropriate standardised criteria were used for appraising primary studies, whether critical appraisal was conducted by two or more

Table 7

Summary of non-systematic review papers

Authors	Type of review	Type of measures	Population	Region or Country	Short summary	Summary of Main Findings
Campbell 2005	Non-systematic	Screening	Children and Adolescents up to the age of 22	NR	Reviews and evaluates screening questionnaires and rating scales that are available for the identification of Asperger's.	Reviews five rating scales used for the screening of ASD. There were consistent limitations in standardisation and norming procedure and all scales showed significant weaknesses. The Krug Asperger's Disorder Index (KADI) was the strongest in terms of reliability and validity.
Charak & Stella 2002	Other review (non-systematic)	Screening and diagnostic	Children, adolescents, and young adults	NR	Reviews measures that are used in the differential diagnosis of ASD. Provides characteristics and information about each measure and the intended use.	Reviews nine measures for ASD. An increasing number psychometrically robust screening and diagnostic measure for ASD have become available. The Autism Diagnostic Observation Schedule (ADOS) demonstrated good reliability and sensitivity as a diagnostic instrument.
Klose et al., 2012	Other review (non-systematic)	Screening and diagnosis	Children, adolescents, and adults	NR	Review ASD measures against the diagnostic criteria provided in the Diagnostic and Statistical Manual of Mental Disorders– Fourth Edition– Text Revision (DSM-IV-TR) and the Individuals with Disabilities Education Act (IDEA) 2004 United States guidelines. Examines the characteristics of these measures and highlights recommendations for school-based assessment teams.	Identifies 5 measures for ASD. The Autism Diagnostic Observation Schedule (ADOS) and Autism Diagnostic Interview-Revised (ADI-R) were found to have the strongest utility related to the IDEA and DSM-IV-TR Criteria. There was a lack of ASD measures for children who speak languages other than English.
Matson et al., 2007	Other Review (Non-systematic)	Screening and diagnostic	Children, adolescents, and adults	NR	Reviews measures for the differential diagnosis of autism, Examines the psychometric properties of these measures and provides directions for future research.	Reviews 21 measures for ASD. There were multiple studies investigating psychometrics for the Autism Behaviour Checklist (ABC), Childhood Autism Rating Scale (CARS), and Autism Diagnostic Observation Schedule (ADOS). The Autism Diagnostic Interview-Revised (ADI-R) has a broader age range of norms and more published psychometric data than other measures. The largest limitation is the is the lack of adequate norms for most measures.
Morgan 1988	Other review (non-systematic)	Diagnostic	Children and adolescents	NR	Evaluates 5 objective scales for diagnosis of autism with reference to psychometric criteria of reliability and validity.	Assesses five diagnostic scales for ASD. The Childhood Autism Rating Scale (CARS) was the strongest scale in terms of demonstrated psychometric properties, with

						acceptable reliability and internal consistency. The other scales showed inconsistent psychometric qualities and there was a lack of data.
Norris & Lecavalier 2010	Other review (non-systematic)	Screening	Child and adolescent	NR	Reviews caregiver-rated screening measures for ASD in children and adolescents above the age of 3 years.	Reviews five caregiver rating scales for ASD. The Social Communication Questionnaire (SCQ) was the strongest in terms of psychometrics and was recommended. The Social Responsiveness Scale (SRS) and Autism Spectrum Screening Questionnaire (ASSQ) showed promise, but further research is needed on these scales.
Parks 1983	Other review (non-systematic)	Screening and diagnostic	Children and adolescents		Examines 5 measures specifically designed to assess autism and discusses reliability and validity issues for these measures.	Reviews five measures for autism. Reliability was acceptable for all scales except Rimland's Diagnostic Checklist. All scales had a lack of discriminant and/ or content validity. Further research examining psychometric properties is needed.
Sappok et al., 2015	Other review (non-systematic)	Screening	Children, adolescents, and adults		Reviews ASD screening measures that are available in English and German. Examines the target populations and characteristics of these ASD screening measures.	Identified 46 screening measures for ASD. Most were designed for children, and a few were designed for adults. The Social Communication Questionnaire (SCQ) has been examined in various populations and the Childhood Autism Rating Scale (CARS), Social Responsiveness Scale (SRS), Autism Spectrum Quotient (AQ) and Empathy Quotient (EQ) were the most widely studied. Research on the other measures was limited.
Stoesz et al., 2011	Other review (non-systematic)	Screening	Adults 18 + years	NR	Identifies five instruments designed specifically for identifying Asperger's Disorder (AD) in adults; describes their documented psychometric properties; and evaluates their utility for the assessment of AD in adult populations.	Evaluates five measures for AD. There is limited normative information that is provided. Evidence of the reliability and validity for each measure was relatively poor. Further research, particularly in the areas of reliability and validity, is needed for each measure.
Thabtah & Peebles 2019	Other (non-systematic review)	Screening	Children, adolescents, and adults	NR	Reviews 37 different ASD screening measures and examines the characteristics and psychometric properties of these measures. Attempts to identify possible areas in future need of development and innovation.	Evaluates 37 screening tools for ASD. None of the screening tools were found to perform well against all of the parameters. The Autism Spectrum Quotient (AQ-10) can be recommended for adolescents and adults. The findings highlight the need for a more efficient screening tool that performs well.

NR: Not Reported

Table 8

Critical appraisal of systematic review papers using JBI checklist

Paper	Is the review question clearly and explicitly stated?	Were the inclusion criteria appropriate for the review question?	Was the search strategy appropriate?	Were the sources and resources used to search for studies adequate?	Were the criteria for appraising studies appropriate?	Was critical appraisal conducted by two or more reviewers independently?	Were there methods to minimize errors in data extraction?	Were the methods used to combine studies appropriate?	Was the likelihood of publication bias assessed?	Were recommendations for policy and/or practice supported by the reported data?	Were the specific directives for new research appropriate?	Total Number of Questions scoring 'Yes' (out of 10)
Backes et al., 2014	✓	?	✓	✓	×	×	✓	✓	✓	✓	✓	8
Baghdadli et al., 2017	✓	✓	✓	✓	✓	?	✓	✓	?	?	✓	8
Falkmer et al., 2013	?	✓	✓	✓	×	×	?	?	?	✓	✓	5
Hirota et al., 2018	?	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	9
Levy et al., 2020	✓	✓	?	✓	✓	✓	✓	✓	✓	✓	✓	10
Soto et al., 2015	✓	✓	?	✓	×	✓	?	✓	×	✓	✓	7
Stewart & Lee 2017	✓	✓	✓	✓	×	×	?	✓	×	✓	✓	7
Sun et al., 2013	✓	✓	✓	✓	×	×	×	✓	✓	✓	?	7
Wang et al., 2020	✓	✓	✓	✓	×	×	?	✓	×	✓	✓	7
Wigham et al., 2019	✓	✓	✓	✓	✓	✓	?	✓	×	✓	✓	9

✓: *Yes*, ×: *No*, ?: *Unsure*

researchers independently, whether there were methods to minimize errors in data extraction and whether the likelihood of publication bias was assessed.

Table 9 shows the quality rating for the 10 non-systematic review papers according to the SANRA scale. Quality ratings were eight and above for most of the non-systematic review papers out of a total sum score of 12. Three of the non-systematic reviews scored 10 or above (Norris & Lecavalier 2010; Sappok et al., 2015; Stoesz et al., 2011) and just two of the non-systematic reviews had a lower score of seven (Charak & Stella 2002; Morgan 1988). There were concerns around the description of the literature search for many of the non-systematic reviews, with just two of the non-systematic review papers describing the literature search in detail (Norris & Lecavalier 2010; Sappok et al., 2015). There were also issues highlighted with the statement of concrete aims and formulation of questions in the non-systematic reviews.

Overlap of the Primary Studies

Analysis of overlap was not conducted for any reviews that did not have a complete list of included primary studies (Pieper et al., 2014). Therefore, the Corrected Covered Area (CCA) was only calculated across the 10 systematic reviews. The overall CCA across the 10 systematic reviews was 0.0129 (1.3%), which showed that there was only slight overlap in the primary studies included by the reviews according to the thresholds laid out by Pieper et al. (2014). The overall CCA matrix was examined, and the CCA was then calculated between any two systematic reviews that had included more than two of the same primary studies. Only two of the systematic review studies had a very high degree of overlap. The CCA between Bagdadli et al., (2017) and Wigham et al., (2019) was 0.16 (16%). Table 10 presents the full CCA calculations.

Table 9

Critical appraisal of non-systematic review papers using the SANRA scale

Paper	Justification of the article's importance for the readership	Statement of concrete aims or formulation of questions	Description of the literature search	Referencing	Scientific reasoning	Appropriate presentation of data	Total (out of 12)
Campbell 2005	1	2	1	1	2	2	**9**
Charak & Stella 2002	1	1	0	2	1	2	**7**
Klose et al., 2012	2	1	0	2	1	2	**8**
Matson et al., 2007	2	1	0	2	2	1	**8**
Morgan 1988	2	1	0	2	1	1	**7**
Norris & Lecavalier 2010	2	1	2	2	2	2	**11**
Parks 1983	2	1	0	2	2	2	**9**
Sappok et al., 2015	2	1	2	2	1	2	**10**
Stoesz et al., 2011	2	2	0	2	2	2	**10**
Thabtah & Peebles 2019	2	2	0	2	1	1	**8**

0 = Not at all, 1 = partially addressed and 2 = fully addressed

Table 10

Calculations of the Corrected Covered Area (CCA) showing overlap of the included primary studies

Review Papers	**CCA**	**Degree of overlap according to (Pieper et al., 2014)***
Bagdadli et al., 2017 & Wigham et al., 2019	0.160 **(16%)**	Very High
Backes et al., 2014 & Soto et al., 2015	0.069 **(6.9%)**	Moderate
Soto et al., 2015 & Stuart & Lee 2017	0.065 **(6.5%)**	Moderate
Bagdadli et al., 2017 & Hirota et al., 2018	0.040 **(4%)**	Slight
Bagdadli et al., 2017 & Falkmer et al., 2013	0.019 **(1.9%)**	Slight
Overall overlap across 10 included systematic review papers (Backes et al., 2014; Bagdadli et al., 2017; Falkmer et al., 2013; Hirota et al., 2018; Levy et al., 2020; Soto et al., 2015; Stuart & Lee 2017; Sun et al., 2012; Wang et al., 2020; Wigham et al., 2019)	0.013 **(1.3%)**	Slight

*Thresholds for degree of overlap: 0-5% Slight; 6-10% Moderate; 11-15% High; >15% Very High (Pieper et al., 2014)

Summary of Identified Autism Measures

A total of 28 screening measure and 15 diagnostic measures for autism were identified from the included reviews. Assessment measures that were designed for sole use with children under the age of 6 years have not been included in this meta-review. The Behaviour Rating Instrument for Autistic and Atypical Children (BRIAAC; Ruttenberg et al., 1966) measure was not included as it compares the behaviour of children against a normally developing 3.5- to 4.5-year-old. The Waterville Autistic Behaviour Scales (WABS; Song et al., 2009) and the Clancy Autism Behaviour Scale (CABS; Clancy et al., 1969) were also not included as the age range for the measures were not available. In addition to this, measures identified by the included reviews that were defined as solely for use with intellectual disabilities have not been included.

Screening Measures

The characteristics of the 28 screening measures are detailed in Table 11, including the original measure, the authors and year of publication, any updated or short version identified in the review papers, age range, administration time, administration mode and a short summary. The screening measures covered a range of ages, and eight of the measures could be used with children, adolescents, or adults. These were the Autism Behaviour Checklist (ABC), Autism Screening Questionnaire (ASQ), Autism Spectrum Quotient (AQ-50), Autistic Traits Assessment Scale (ATA), Krug Asperger's Disorder Index (KADI), Marburg Rating Scale for Asperger's Syndrome (MBAS) (Only available in German), Social Communication Questionnaire (SCQ) and the Social Responsiveness Scale (SRS). Five of the screening measures were for use with adults only including the Autism Checklist (ACL), Adult Social Behaviour Questionnaire (ASBQ), Autism Spectrum Disorder in Adults Screening questionnaire (ASDASQ), Empathy Quotient (EQ) and the Sensory Reactivity in Autism Spectrum (SR-AS).

The remaining screening measures were designed for use with children and adolescents. Six of the screening measure were designed for use with children up to the age of 13 and these included the Australian Scale for Asperger's Syndrome (ASAS), Autism Screening Instrument for Educational Planning (ASIEP), Autism Symptom Interview (ASI), Childhood Asperger Syndrome Test (CAST), Pervasive Developmental Disorder Behavior Inventory (PDDBI) and the Pervasive Developmental Disorders Rating Scale (PDDRS). There were eight screening measure that had an age range covering children and adolescents which were the Autism Spectrum Rating Scale (ASRS), Autism Spectrum Screening Questionnaire (ASSQ), Child Behaviour Checklist (CBCL), Children's Communication Checklist (CCC), Children's Social Behaviour Questionnaire (CSBQ), Fremdbeurteilungsfragebogen für tiefgreifende Entwicklungsstörungen (FBB-TES) (only

Table 11

Summary of autism screening measures

Name of Measure	Short or Updated Versions	Review Papers identified in	Authors and Year of Publication	Age Group	Administration Time	Administration Mode	Short Summary
Autism Behaviour Checklist (ABC)		Backes et al., 2014 Hirota et al., 2018 Stuart & Lee 2017 Sun et al., 2012 Wang et al., 2020 Charak & Stella 2002 Matson et al., 2007 Parks 1983 Sappok et al., 2015 Thabtab & Peebles 2019	Krug, Arick, & Almond 1980	3 years and older	15 minutes	Teacher or healthcare questionnaire	The ABC is one part of the Autism Screening Instrument for Educational Planning (ASIEP). The ABC is a 57-item behaviour rating scale that consists of five categories: sensory, relating, body and object use, language and social and self-help.
Autism Checklist (ACL)		Wigham et al., 2019 Sappok et al., 2015	Sipes & Matson 2014	Adults	10 minutes	Clinician observational measure	The ACL assesses the domains of social interaction, social communication, and stereotyped and restrictive behaviours. Each domain has 4 items that are scored as present, partly present or not present. Official training is not require but administration should be conducted by clinicians with 'ASD expertise'.
Autism Spectrum Quotient (AQ-50) Original version		Baghdadli et al., 2017 Hirota et al., 2018 Soto et al., 2015	Baron-Cohen, Wheelwright, Skinner, Martin & Clubley 2001	4 years - Adults	5-10 minutes	Self-questionnaires	The AQ is a self-report questionnaire designed to measure ASD traits in an individual by their own assessment.

	The Short Autism Spectrum Quotient (AQ-10) Autism Quotient-Short form (AQ-S or AQ28) Autism Quotient–20 (AQ-20) Autism Quotient Japanese version–21 (AQ-J-21) Autism Quotient–39 (AQ-39)	Stuart & Lee 2017 Wigham et al., 2019 Sappok et al., 2015 Thabtab & Peebles 2019	Allison, Auyeung, Baron-Cohen 2012 Hoesktra et al. 2011 Brugha et al. 2012 Kurita, Koyama & Osada 2005 Lau, Kelly & Peterson 2013				
Adult Social Behaviour Questionnaire (ASBQ)		Baghdadli et al., 2017	Horwitz 2016	Adults	NR	Either self-report or informant-report questionnaires	The ASBQ is a 44 item questionnaire that has six domains: reduced contact, reduced empathy, reduced interpersonal insight and theory of mind, violation of social conventions, insistence on sameness, sensory stimulation & motor stereotypes.
Autism Spectrum Disorder in Adults Screening questionnaire (ASDASQ)		Baghdadli et al., 2017 Sappok et al., 2015 Stuart & Lee 2017	Nylander & Gillberg 2001	Adults	NR	Clinician rating scale	The ASDASQ is a clinician- rated Questionnaire that is completed using observations of the client's behaviour. There are nine questions about symptom/impairments and one question regarding previous contact with services.

Australian Scale for Asperger's Syndrome (ASAS)		Sappok et al., 2015 Thabtab & Peebles 2019	Garnett & Attwood 1995	Primary school-aged children 6-12 years	10-15 minutes	Parental questionnaires administered by a clinician	The ASAS is a 25-item questionnaire that is administered by a clinician to identify Asperger's behaviours. There are five subscales: social and emotional abilities, communication skills, cognitive skills, specific interests, and motor skills.
Autism Symptom Interview (ASI)		Hirota et al., 2018	Bishop et al., 2017	5 to 12 year old children	15-20 minutes	Phone interview administered by examiners	The ASI is a short phone interview that based on ADI-R questions. It can be delivered by examiners with little training.
Autism Screening Instrument for Educational Planning (ASIEP)	ASIEP-2 ASIEP-3	Morgan 1988 Thabtab & Peebles 2019	Krug, Arick & Almond 1980	2-13 years	Varies	Questionnaires and activities	The ASIEP is used to evaluate children with autism and can be used for differential diagnosis. It is used to develop educational plans and monitor progress. The ASIEP-3 is composed of the following subtests: Autism Behavior Checklist, Sample of Vocal Behavior, Interaction Assessment, Educational Assessment, Prognosis of Learning Rate.
Autism Screening Questionnaire (ASQ)		Backes et al., 2014 Soto et al., 2015 Stuart & Lee 2017 Charak & Stella 2002	Berument et al., 1999	Over 4 years old	Less than 10 minutes	Parent or caregiver questionnaire	The ASQ is a 40-item questionnaire based on the revised version of the Autism Diagnostic Interview (ADI-R).
Autism Spectrum Rating Scale (ASRS)		Wang et al., 2020	Goldstein & Naglieri, 2009	2-18 years	20 minutes	Multi-informant questionnaire	The ASRS is a multi-informant measure that helps to identify ASD through assessment of symptoms and behaviours.
Autism Symptom Self-Report for Adolescents and Adults (ASSERT)		Sappok et al., 2015	Posserud et al. 2013	16-19 years	NR	Self-report online questionnaire	The ASSERT is a self-report 7-item online questionnaire for use with adolescents.

Autism Spectrum Screening Questionnaire (ASSQ)	Autism Spectrum Screening Questionnaire—Mandarin Chinese Version (ASSQ-CV)	Hirota et al., 2018 Soto et al., 2015 Stuart & Lee 2017 Sun et al., 2012 Campbell 2005 Norris & Lecavalier 2010 Sappok et al., 2015 Thabtab & Peebles 2019	Ehlers, Gillberg, & Wing, 1999	5-18 years	5-10 minutes	Parent and teacher rated questionnaire	The ASSQ is a 27-item questionnaire with 4 domains: social interaction, communication problems, restricted and repetitive behaviour, and motor clumsiness and other associated symptoms. Items are rated on a 3 point scale.
Autistic Traits Assessment Scale (ATA)		Backes et al., 2014	Ballabriga et al. 1994	Over 2 years old	From 20 to 30 minutes	Questionnaire based on direct observation	The ATA is composed of 23 subscales that are coded based on based on direct observation of the child.
Childhood Asperger Syndrome Test (CAST)		Hirota et al., 2018 Stuart & Lee 2017 Wang et al., 2020 Campbell 2005 Sappok et al., 2015 Thabtab & Peebles 2019	Scott, Baron-Cohen, Bolton, & Brayne, 2002	4-11 years	5-10 minutes	Parent-report questionnaire	The CAST is a 37-item, parent-report questionnaire of behaviours relating to Asperger's. Items are scored as present or absent. A cut-off score of 15 or more indicates the need for further testing.
Child Behaviour Checklist (CBCL)		Stuart & Lee 2017 Thabtab & Peebles 2019	Bordin et al., 2013	6-18 years	NR	Parent-report questionnaire	The CBCL is a 118-item parent-report questionnaire. It consists of 8 domains: Attention problems, aggressive behaviour, anxiety levels, rule breaking behaviour, social problems, somatic complaints, depression levels, and thought problems. Each domain represents one subscale on the CBCL.

Children's Communication Checklist (CCC)	Hirota et al., 2018	Bishop, 1998	4-16 years old	5-15 minutes	Rating scale	The CCC is a 70-item rating scale that assesses three areas: language structure, autistic behaviour and pragmatic communication.
Children's Social Behaviour Questionnaire (CSBQ)	Sappok et al., 2015	Luteijn et al. 2000	4-18 years	NR	Parent-rating questionnaire	The CSBQ is a 49-item parent-report questionnaire. The questionnaire quantifies different behavioural dimensions associated with ASD. The CSBQ assesses six behavioural dimensions.
Empathy Quotient (EQ)	Sappok et al., 2015	Baron-Cohen & Wheelwright 2004	Adults	NR	Self-rating questionnaire	The Empathy Quotient (EQ) is a 60-item questionnaire. The EQ measures empathy in adults.
Fremdbeurteilungsfragebogen für tiefgreifende Entwicklungsstörungen (FBB-TES) ***Only available in German***	Sappok et al., 2015	Döpfner & Lehmkuhl 1998	Children	NR	Close carer-rated scale	The FBB-TES is a 14-item carer-rated questionnaire that is based on the ICD-10/DSM-IV. The FBB-TES is part of the DISYPS-KJ (Diagnostiksystem für psychische störungen im kindes– und jugendalter).
Krug Asperger's Disorder Index (KADI)	Campbell 2005 Sappok et al., 2015 Stoenz et al., 2011 Thabtab & Peebles 2019	Krug 2003	6-21 years	15-20 minutes	Parent or teacher-rated questionnaire	The KADI is a 32-item screening questionnaire used to identify Asperger's disorder.
The Marburg Rating Scale for Asperger's Syndrome (MBAS)	Sappok et al., 2015	Kamp-Becker et al. 2005	6-24 years	NR	Informant or close carer rated questionnaire	The MBAS is a 70-item questionnaire based on the criteria for Asperger syndrome in the ICD-10/DSM-IV.
Pervasive Developmental Disorder Behaviour Inventory (PDDBI)	Falkmer et al., 2013 Matson et al., 2007	Cohen et al., 2003	18 months - 12 years	NR	Informant-rating questionnaire	The PDDBI is an informant-based rating scale that measures the effectiveness of treatments for children with developmental disabilities, including ASD. The

						areas it assesses are behavioural challenges, key skills and abilities.
Pervasive Developmental Disorders Rating Scale (PDDRS)	Matson et al., 2007	Eaves, Campbell & Chambers 2000	1-12 years	20 minutes	Parent or teacher-rating scale	The PDDRS is a screening instrument that measures 3 factors of arousal, affect and cognition and is used to identify individuals with ASD and other pervasive developmental disorders.
Social and Communication Disorders Checklist (SCDC)	Hirota et al., 2018 Soto et al., 2015 Stuart & Lee 2017 Sappok et al., 2015	Skuse, Mandy, & Scourfield, 2005	Children and adolescents	NR	Parent-rating questionnaire	The SCDC is a 12-item parent rating questionnaire that measures 4 areas: reciprocal social interaction skills, communication skills and general behavioural problems and functional impairment. There is a cut-off score of 9 for identifying potential ASD.
Social Communication Questionnaire (SCQ)	Hirota et al., 2018 Levy et al., 2020 Soto et al., 2015 Stuart & Lee 2017 Wang et al., 2020 Wigham et al., 2019 Norris & Lecavalier 2010 Sappok et al., 2015 Thabtab & Peebles 2019	Rutter, Bailey, & Lord, 2003 Derks et al., 2017	4 years and up	10 minutes	Parent-report questionnaire	The SCQ is a 40-item parent-report questionnaire about autistic behaviour. The cut-off score is 15 or above.
Sensory Reactivity in Autism Spectrum (SR-AS)	Baghdadli et al., 2017	Elwin, Schroder, Ek & Kjellin 2015	Adults	NR	Self-questionnaire	The SR-AS is a self-report questionnaire that assesses 4 areas: Awareness/Hyper-reactivity, Low Awareness/Hypo-reactivity and Sensory interest and motor. Higher

							scores indicate higher frequencies of sensory reactivity.
Social Responsiveness Scale (SRS)	Social Responsiveness Scale 2nd edition-Adult Form (SRS-Adult or SRS-A) SRS2-AS30 SRS2-AS11	Baghdadli et al., 2017 Hirota et al., 2018 Soto et al., 2015 Stuart & Lee 2017 Wang et al., 2020 Wigham et al., 2019 Norris & Lecavalier 2010 Sappok et al., 2015 Thabtab & Peebles 2019	Constantino & Gruber 2012 Duku et al. 2013 Kanne et al. 2009 Reiersen et al. 2008	4 years and up	15-20 minutes	Questionnaire	The SRS is a 65-item parent/teacher report questionnaire that assesses the presence and severity of ASD traits and social impairment. The cut off score for ASD is a T score of 60 or higher.
Wing Subgroup Questionnaire (WSQ)		Sappok et al., 2015	Castelloe & Dawson 1993	Children (6-19 years)	NR	Carer-report questionnaire	The WSQ is a 50-items carer-report questionnaire with 3 of “aloof”, “passive and friendly”, “active-but-odd”.

NR: Not Recorded

available in German), Social and Communication Disorders Checklist (SCDC) and the Wing Subgroup Questionnaire (WSQ). The FBB-TES specified did not specify an exact age range but was described as for use with children. There was one screening measure, the Autism Symptom Self-Report for Adolescents and Adults (ASSERT) that was designed for use with just adolescents between the ages of 16 and 19.

Twenty-one of the screening measures were parent-, teacher-, clinician- or self-report questionnaires, with an administration time of between 5-20 minutes. There were two observation rated scales, the ACL and ATA which required direct observation of behaviour. There was one screening measure which was administered as a phone interview by a trained examiner, the ASI, and one screening measure which was a combination of questionnaires and activities, the ASIEP. However, these screening measures that had different administration modes only had slightly longer administration times between 10-30 minutes.

Diagnostic Measures

The characteristics of the 15 diagnostic measures are detailed in Table 12 including the original measure, the authors and year of publication, any updated or short version identified in the review papers, age range, administration time, administration mode and a short summary. Six of the measures included here were categorised as both diagnostic and screening by different review papers.

Nine of the diagnostic measures were designed for use with children, adolescents, or adults. These were the Autism Diagnostic Observation Schedule (ADOS), Autism Diagnostic Interview (ADI), Autism Mental Status Exam (AMSE), Asperger Syndrome Diagnostic Interview (ASDI), Childhood Autism Rating Scale (CARS), Diagnostic Interview for Social and Communication Disorder (DISCO), Gilliam Asperger's Disorder Scale (GADS), Gilliam Autism Rating Scale (GARS) and the Developmental, dimensional and diagnostic Interview

Table 12

Summary of autism diagnostic measures

Name of Measure	Short or Updated Versions	Review Papers included in	Authors and Year of Publication	Age Group	Administration Time	Administration Mode	Short Summary
Adult Asperger Assessment (AAA)*		Baghdadli et al., 2017 Falkmer et al., 2013 Sappok et al., 2015 Stoenz et al., 2011	Baron Cohen, Wheelwright, Robinson & Woodbury-Smith 2005	Adults with average IQ minimum	Around 3 hours	Self-report questionnaires and guide to clinical interview with client and/or informant	The AAA is based on self-rating in AQ and EQ and then an expert-interview which is based on the criteria for ASD.
Autism Diagnostic Interview (ADI)	Autism Diagnostic Interview–Revised (ADI-R)	Backes et al., 2014 Falkmer et al., 2013 Sun et al., 2012 Wigham et al., 2019 Klose et al., 2012 Matson et al., 2007	LeCouteur et al., 1989 Lord, Rutter and LeCouteur 1994	Over 2 years old	From 1 hour 30 minutes to 2 hours 30 minutes	Interview	The ADI is a standardized, semi-structured interview conducted with a caregiver that assess a range of behaviours associated with ASD.
Autism Diagnostic Observation Schedule Generic (ADOS-G)	Second Edition (ADOS-2)	Baghdadli et al., 2017 Falkmer et al., 2013 Sun et al., 2012 Wigham et al., 2019 Charak & Stella 2002 Klose et al., 2012	Lord, Rutter, DiLavore, & Risi 2000 Lord, Rutter, et al., 2012	Over 12 months	30-60 minutes	Semi-structured, standardised observational assessment	The ADOS is a standardised semi-structured diagnostic assessment. The ADOS consists of one-to-one interaction and direct observation of an individual by a trained examiner. A range of semi-structured activities are used to assess for ASD.

	Matson et al., 2007					
Autism Mental Status Exam (AMSE)	Baghdadli et al., 2017 Wigham et al., 2019 Sappok et al., 2015	Grodberg et al., 2014	18 months - adults		Observation clinician rating scale	The AMSE is a clinician-rated 8-item scale that measures social, communicative, and behavioural functioning. It utilises direct clinical observation and parent report.
Autism Spectrum Disorders-Diagnosis for Children (ASD-DC)*	Falkmer et al., 2013 Sappok et al., 2015 Matson et al., 2007	Matson, Terlonge, & Gonzalez 2006	2-16 years	30-45 minutes	Trained examiner or parent rating scale	The ASD-DC is a 40-item scale that is scored in a 3-point Likert format: 0 (not different; no impairment), 1 (somewhat different; mild impairment), or 2 (very different; severe impairment). The ASD-DC is used to measure autism, PDD-NOS, and Asperger's Syndrome.
Asperger Syndrome Diagnostic Interview (ASDI)	Falkmer et al., 2013 Sappok et al., 2015 Stoenz et al., 2011	Gillberg et al., 2001	Children and adults	NR	Standardised interview	The ASDI is a short, standardized interview compromised of 20 binary items and based on Gillberg's criteria. There are 6 domains.
Asperger Syndrome Diagnostic Scale (ASDS)*	Falkmer et al., 2013 Campbell 2005 Charak & Stella 2002 Norris & Lecavalier 2010 Sappok et al., 2015 Thabtab & Peebles 2019	Myles, Bock & Simpson 2001	5-18 years	10-15 minutes	Parent, teacher or clinician rated questionnaire	The ASDS is a 50-item parent, teacher or clinician rated questionnaire consists of five subscales: language, social, maladaptive, cognitive, and sensorimotor.
Childhood Autism Rating Scale (CARS)*	Backes et al., 2014 Stuart & Lee 2017	Schopler et al., 1980	Over 2 years old	From 5 to 10 minutes	Questionnaire based on direct observation	The CARS is a 15-item behavioural rating scale that is used to assess ASD.

	Childhood Autism Rating Scale 2nd Editions (CARS-2)	Sun et al., 2012 Charak & Stella 2002 Klose et al., 2012 Matson et al., 2007 Morgan 1988 Parks 1983 Sappok et al., 2015 Thabtab & Peebles 2019					
Checklist for Autism Spectrum Disorder (CASD)		Falkmer et al., 2013	Mayes 2001	1-17 years	15 minutes	Questionnaire	The CASD is a questionnaire that rates 30 symptoms of ASD across six areas: Problems with social interaction, Perseveration, Somatosensory disturbance, atypical communication and development, Mood and Problems with attention and safety
Developmental, dimensional, and diagnostic Interview (3Di)		Falkmer et al., 2013 Matson et al., 2007	Skuse et al., 2004	3 years and up	Around 50 minutes	Structured interview by trained examiner	The 3Di is a standardized interview that is computer-based. The 3Di assesses score in terms of their severity, frequency, and comorbidity related to ASD.
Diagnostic Interview for Social and Communication Disorder (DISCO)		Falkmer et al., 2013 Matson et al., 2007	Wing et al., 2002	Children and adults	NR	Interviewer-based schedule for use with parents and carers	The DISCO is an interview-based schedule for the diagnosis of ASD and helps to assess individual needs. Information is recorded for a wide range of behaviours and developmental skills.
Gilliam Asperger's Disorder Scale (GADS)		Falkmer et al., 2013 Campbell 2005 Sappok et al., 2015 Stoenz et al., 2011	Gilliam 2001	3-22 years	10 minutes	Parent, teacher or clinician rated scale	The GADS is a 32-item questionnaire used to assess Asperger's Disorder. The GADS includes four subscales: Social Interaction, Restricted Patterns, Cognitive Patterns, and Pragmatic Skills.

		Thabtab & Peebles 2019					
Gilliam Autism Rating Scale (GARS)*	Gilliam Autism Rating Scale 2nd Edition (GARS-2)	Falkmer et al., 2013 Stuart & Lee 2017 Charak & Stella 2002 Klose et al., 2012 Matson et al., 2007 Norris & Lecavalier 2010 Sappok et al., 2015	Gilliam 1995	3-22 years	5-10 minutes	Parent/caregiver and teacher rating scale	The Gilliam Autism Rating Scale is a 42-item standardized measure. It is used to assess and diagnose autism and other behavioural conditions. It utilises parental or teacher reports of the child's behaviour and interaction and has a short administration time.
Rimland Diagnostic Checklist, Form E2 (E2-DC)		Matson et al., 2007 Morgan 1988 Parks 1983	Rimland 1971	Children	NR	Parent or teacher-rating scale	The E2-DC is a parent or teacher-report checklist to help with the diagnosis of ASD. It also asks parents to rate the effectiveness of any interventions that have been tried.
Ritvo Asperger and Autism Diagnostic Scale-Revised (RAADS-R)*	RAADS-14 screen	Baghdadli et al., 2017 Hirota et al., 2018 Sappok et al., 2015 Stoenz et al., 2011 Thabtab & Peebles 2019	Ritvo et al., 2011 Erikkson, Andersen & Bejerot 2013	Adults	NR	Clinician administered self-questionnaire	The RAADS-R is an 80-item clinician-administered questionnaire. The RAADS-R is used with adults and is used to identify ASD.

**Sometimes categorised as screening and sometimes described as diagnostic; NR: Not Recorded*

(3Di). Four of the diagnostic measures were specified for use with children and adolescents which were the Autism Spectrum Disorders-Diagnosis for Children (ASD-DC), Asperger Syndrome Diagnostic Scale (ASDS), Checklist for Autism Spectrum Disorder (CASD) and the Rimland Diagnostic Checklist (E2-DC). There were just two diagnostic measures that were designed for use with adults only: the Adult Asperger Assessment (AAA) and the Ritvo Asperger and Autism Diagnostic Scale-Revised (RAADS-R).

There was more variety in the mode of administration in the diagnostic measures. Seven of the diagnostic measures were administered as questionnaires that were either self-, parent-, teacher- or clinician- rated (ASD-DC; ASDS; CASD; E2-DC; GADS; GARS; RAADS-R), with one questionnaire being observation based (CARS). Five of the diagnostic based measures were interview-based measures delivered by trained examiners to parents, caregivers, or the individual with standardised or structured questions (AAA; ADI; ASDI; DISCO; 3Di). These diagnostic measures had a longer administration time from 30-180 minutes. One of the diagnostic measures, the ADOS, was administered differently as a semi-structured, standardised observational assessment with a trained clinician. This also had a longer administration time of between 30-60 minutes.

Psychometric Properties of the Measures

Harvest plots were produced for each screening and diagnostic measure to visually represent information about the psychometric properties that was present in the included review papers. These harvest plots were then combined into a matrix representing the psychometric properties of the screening measures (Figure 2) and a matrix representing the psychometric properties of the diagnostic measures (Figure 3). The harvest plots include details about internal consistency, reliability, validity, sensitivity, and specificity synthesised from the psychometric information in the included review studies. Due to time and resource constraints, it was not pragmatically possible to go into each of the individual primary studies

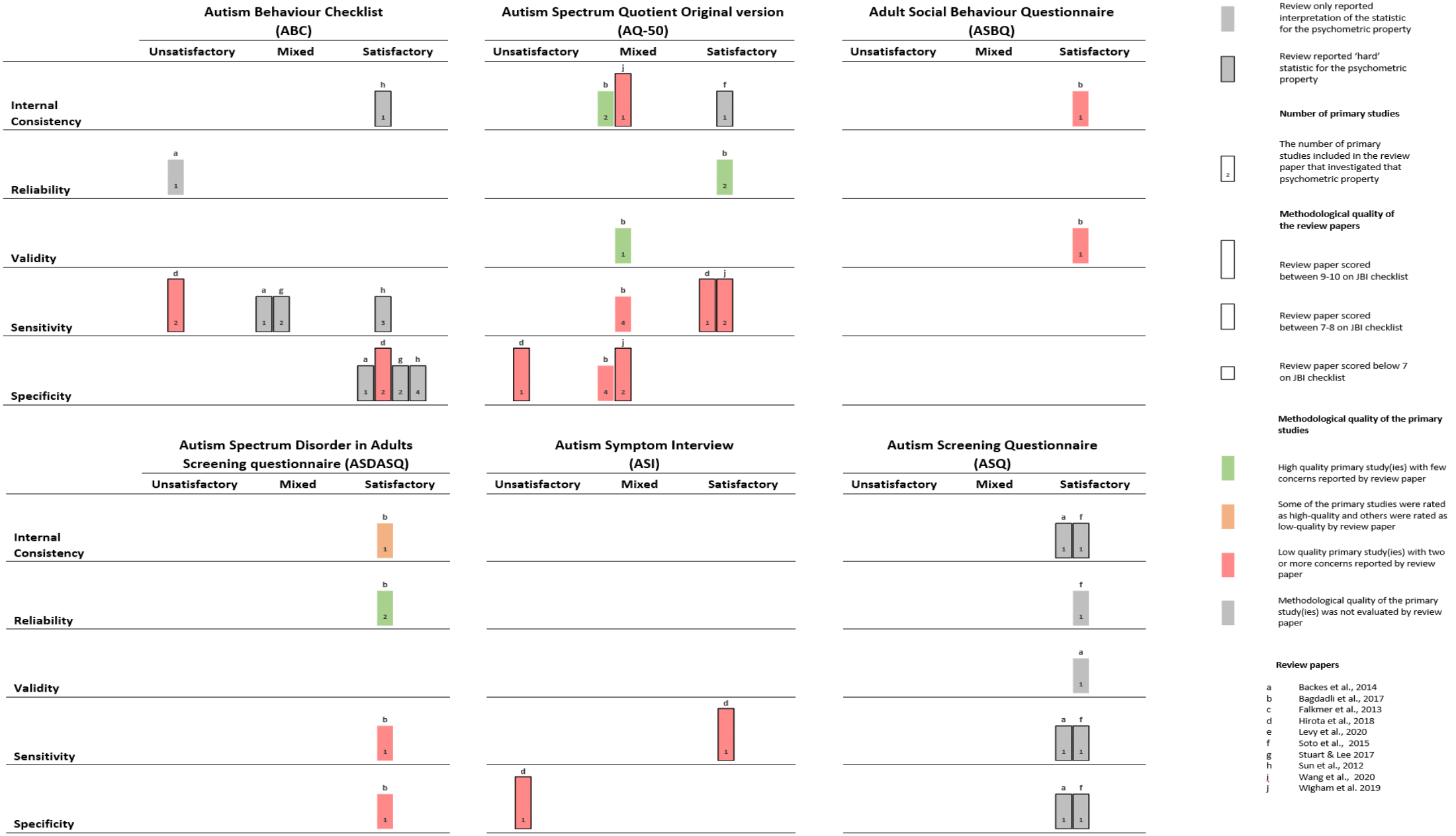
to gather psychometric data or to assess the quality of the primary studies. Psychometric data could only be synthesised from review papers with a complete list of included primary studies, and therefore psychometric data was not collated from any of the non-systematic review papers.

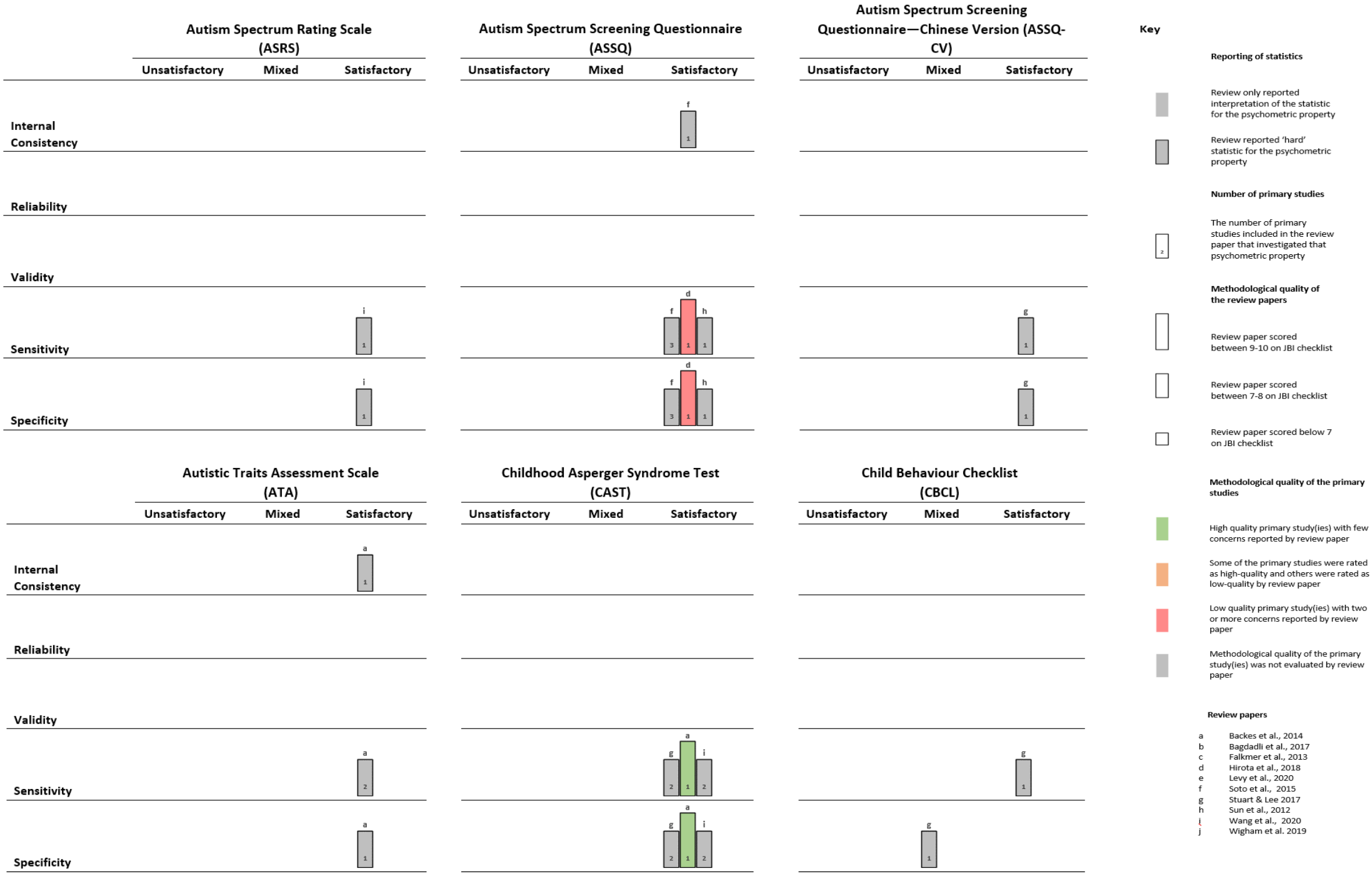
Each harvest plot consists of five rows (one for each type of psychometric property) and three columns (which consist of synthesised quality ratings for each psychometric property; *satisfactory*, *mixed*, and *unsatisfactory*). These ratings represent a summary of the psychometric data from all relevant primary studies in each review. In the harvest plot, each review paper is represented by a bar in each row, where the review reported relevant results regarding psychometrics. Reviews that reported 'hard' statistics are indicated with black outline bars, and reviews that reported interpretations of statistics are indicated by bars with no black outline. Where statistics were available, internal consistency and reliability were rated as *satisfactory* if the Cronbach's alpha coefficient or the Kappa coefficient was ≥70 (Cicchetti, 1994). For sensitivity and specificity values ≥70% were rated as *satisfactory* (Furr, 2011; Glascoe 2005). Any statistics that did not reach these thresholds were rated as *unsatisfactory*. Any range of statistics that spanned above and below these thresholds, were rated as *mixed*. Evidence for criterion, construct or content validity was displayed under the term validity, and the reviews did not report any 'hard' statistics for these three areas.

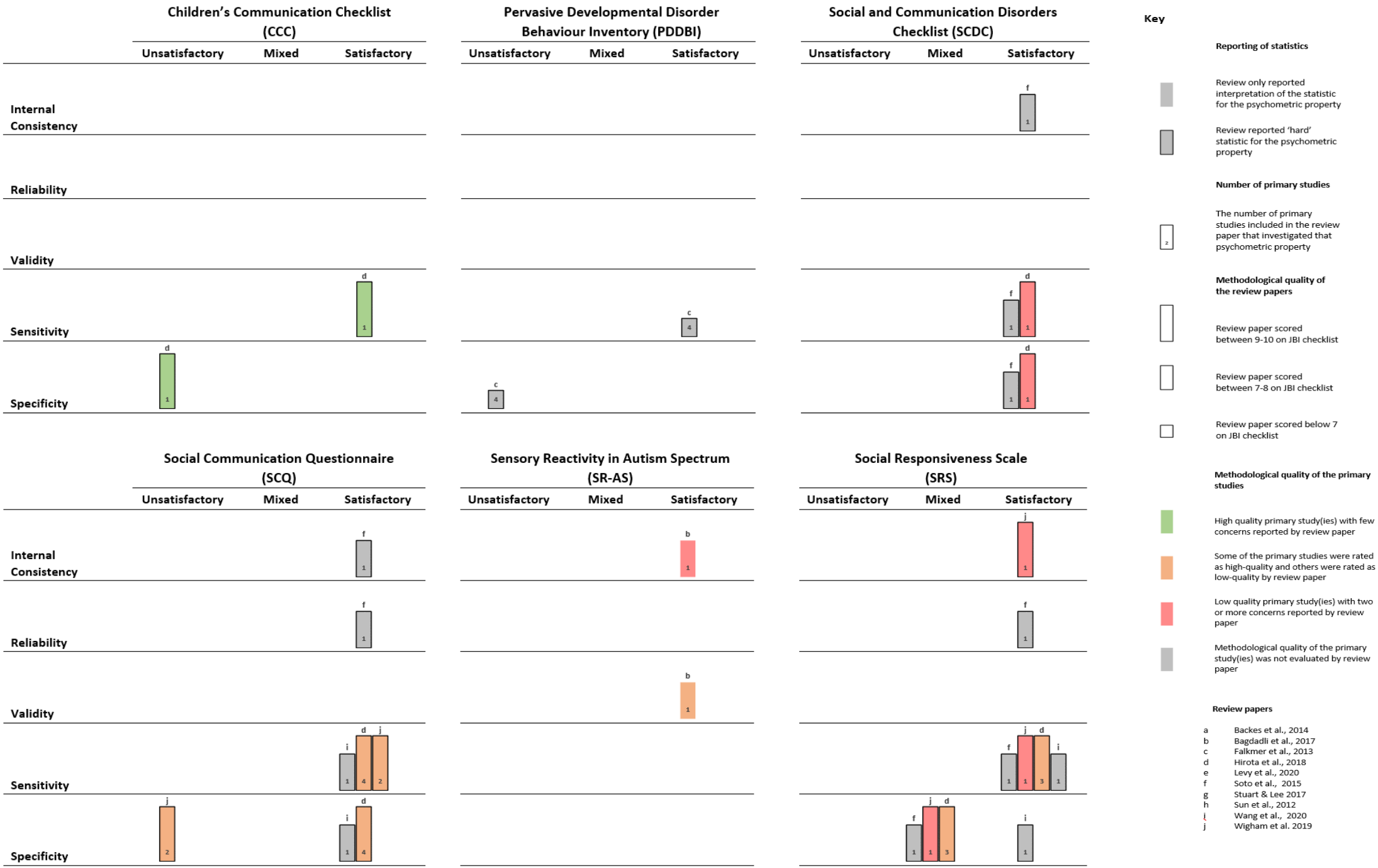
Where the quality of the primary study(ies) was evaluated by the review paper against a standardised tool, this has been indicated by the colour of the bar. The quality of the review paper, critically appraised by this meta-review using the JBI checklist, is indicated by the height of the bar. Each bar is annotated with the number of primary studies included in the review paper, which investigated that psychometric property. The key in Figures 2 and 3 contains more detail.

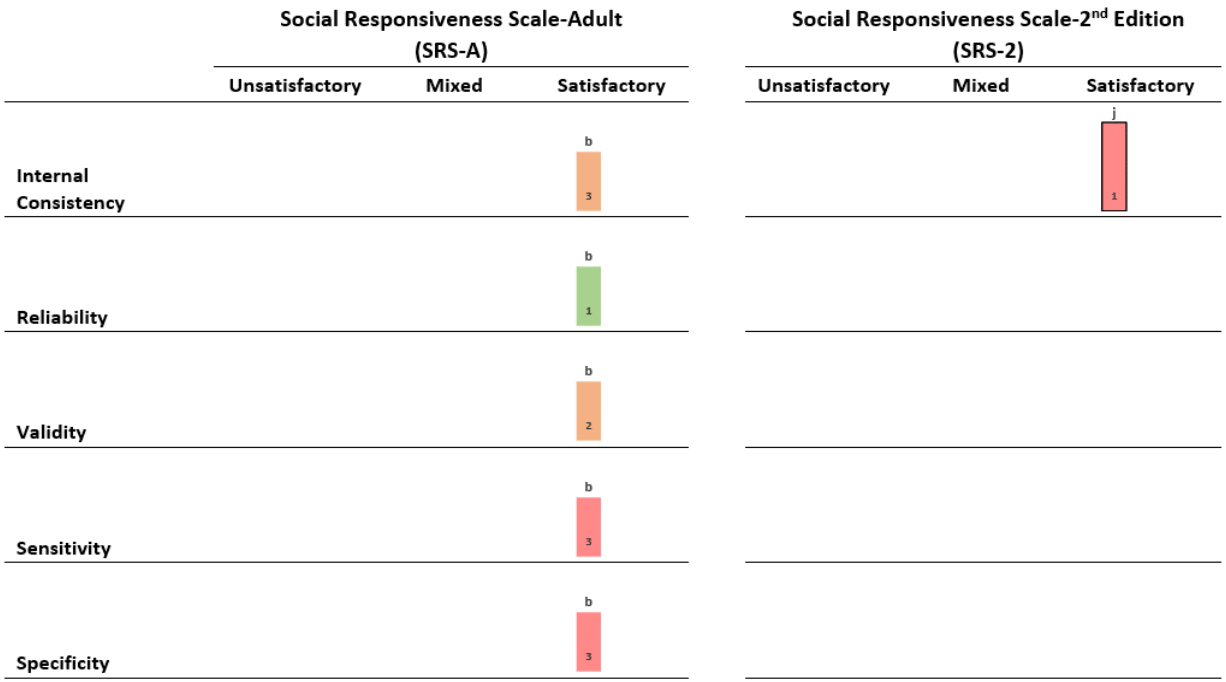
Figure 2

Harvest plot matrix representing psychometric properties for screening measures




Autism Spectrum Rating Scale (ASRS)
Unsatisfactory
Mixed
Satisfactory
Internal Consistency
Reliability
Validity
Sensitivity
Specificity
i
1
Autism Spectrum Screening Questionnaire (ASSQ)
Unsatisfactory
Mixed
Satisfactory
f
1
d
f
h
3
1
1
Autism Spectrum Screening Questionnaire—Chinese Version (ASSQ-CV)
Unsatisfactory
Mixed
Satisfactory
g
1
Autistic Traits Assessment Scale (ATA)
Unsatisfactory
Mixed
Satisfactory
a
1
2
Childhood Asperger Syndrome Test (CAST)
Unsatisfactory
Mixed
Satisfactory
a
g
i
2
1
2
Child Behaviour Checklist (CBCL)
Unsatisfactory
Mixed
Satisfactory
g
1
Key
Reporting of statistics
Review only reported interpretation of the statistic for the psychometric property
Review reported 'hard' statistic for the psychometric property
Number of primary studies
The number of primary studies included in the review paper that investigated that psychometric property
Methodological quality of the review papers
Review paper scored between 9-10 on JBI checklist
Review paper scored between 7-8 on JBI checklist
Review paper scored below 7 on JBI checklist
Methodological quality of the primary studies
High quality primary study(ies) with few concerns reported by review paper
Some of the primary studies were rated as high-quality and others were rated as low-quality by review paper
Low quality primary study(ies) with two or more concerns reported by review paper
Methodological quality of the primary study(ies) was not evaluated by review paper
Review papers
a Backes et al., 2014
b Bagdadli et al., 2017
c Falkmer et al., 2013
d Hirota et al., 2018
e Levy et al., 2020
f Soto et al., 2015
g Stuart & Lee 2017
h Sun et al., 2012
i Wang et al., 2020
j Wigham et al. 2019


Children's Communication Checklist (CCC)
Unsatisfactory
Mixed
Satisfactory
Internal Consistency
Reliability
Validity
Sensitivity
Specificity
Pervasive Developmental Disorder Behaviour Inventory (PDDBI)
Unsatisfactory
Mixed
Satisfactory
Social and Communication Disorders Checklist (SCDC)
Unsatisfactory
Mixed
Satisfactory
Social Communication Questionnaire (SCQ)
Unsatisfactory
Mixed
Satisfactory
Internal Consistency
Reliability
Validity
Sensitivity
Specificity
Sensory Reactivity in Autism Spectrum (SR-AS)
Unsatisfactory
Mixed
Satisfactory
Social Responsiveness Scale (SRS)
Unsatisfactory
Mixed
Satisfactory
Key
Reporting of statistics
Review only reported interpretation of the statistic for the psychometric property
Review reported 'hard' statistic for the psychometric property
Number of primary studies
The number of primary studies included in the review paper that investigated that psychometric property
Methodological quality of the review papers
Review paper scored between 9-10 on JBI checklist
Review paper scored between 7-8 on JBI checklist
Review paper scored below 7 on JBI checklist
Methodological quality of the primary studies
High quality primary study(ies) with few concerns reported by review paper
Some of the primary studies were rated as high-quality and others were rated as low-quality by review paper
Low quality primary study(ies) with two or more concerns reported by review paper
Methodological quality of the primary study(ies) was not evaluated by review paper
Review papers
a Backes et al., 2014
b Bagdadli et al., 2017
c Falkmer et al., 2013
d Hirota et al., 2018
e Levy et al., 2020
f Soto et al., 2015
g Stuart & Lee 2017
h Sun et al., 2012
i Wang et al., 2020
j Wigham et al. 2019



Key

Reporting of statistics

Review only reported interpretation of the statistic for the psychometric property

Review reported 'hard' statistic for the psychometric property

Number of primary studies

The number of primary studies included in the review paper that investigated that psychometric property

Methodological quality of the review papers

Review paper scored between 9-10 on JBI checklist

Review paper scored between 7-8 on JBI checklist

Review paper scored below 7 on JBI checklist

Methodological quality of the primary studies

High quality primary study(ies) with few concerns reported by review paper

Some of the primary studies were rated as high-quality and others were rated as low-quality by review paper

Low quality primary study(ies) with two or more concerns reported by review paper

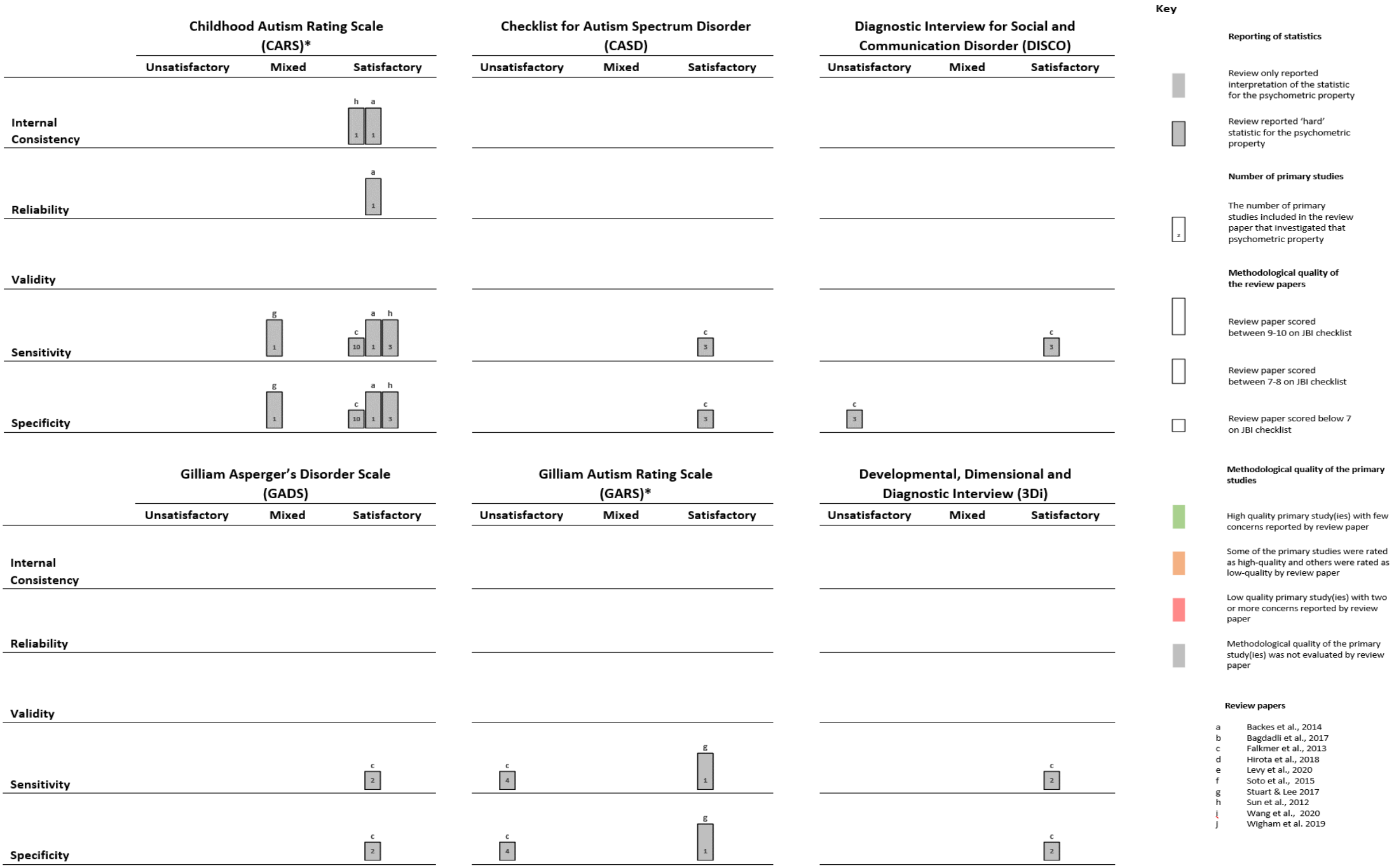
Methodological quality of the primary study(ies) was not evaluated by review paper

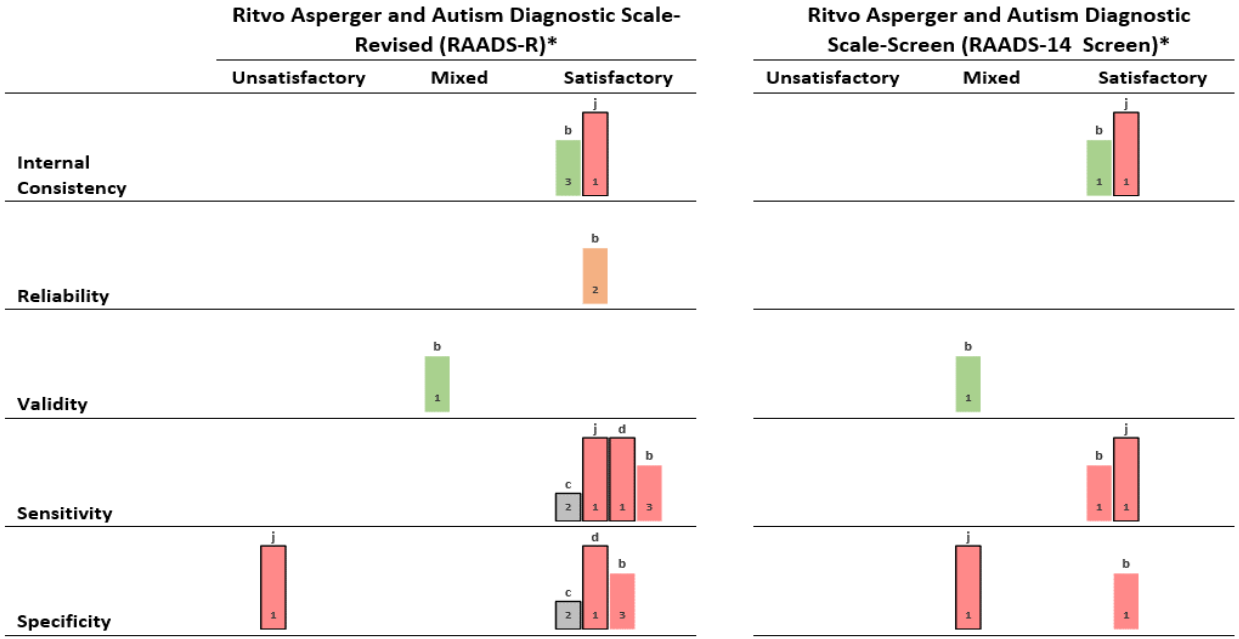
Review papers

a Backes et al., 2014
b Bagdadli et al., 2017
c Falkmer et al., 2013
d Hirota et al., 2018
e Levy et al., 2020
f Soto et al., 2015
g Stuart & Lee 2017
h Sun et al., 2012
i Wang et al., 2020
j Wigham et al. 2019

Figure 3

Harvest plot matrix representing psychometric properties for diagnostic measures


Childhood Autism Rating Scale (CARS)*
Checklist for Autism Spectrum Disorder (CASD)
Diagnostic Interview for Social and Communication Disorder (DISCO)
Unsatisfactory
Mixed
Satisfactory
Internal Consistency
Reliability
Validity
Sensitivity
Specificity
Gilliam Asperger's Disorder Scale (GADS)
Gilliam Autism Rating Scale (GARS)*
Developmental, Dimensional and Diagnostic Interview (3Di)
Key
Reporting of statistics
Review only reported interpretation of the statistic for the psychometric property
Review reported 'hard' statistic for the psychometric property
Number of primary studies
The number of primary studies included in the review paper that investigated that psychometric property
Methodological quality of the review papers
Review paper scored between 9-10 on JBI checklist
Review paper scored between 7-8 on JBI checklist
Review paper scored below 7 on JBI checklist
Methodological quality of the primary studies
High quality primary study(ies) with few concerns reported by review paper
Some of the primary studies were rated as high-quality and others were rated as low-quality by review paper
Low quality primary study(ies) with two or more concerns reported by review paper
Methodological quality of the primary study(ies) was not evaluated by review paper
Review papers
a Backes et al., 2014
b Bagdadli et al., 2017
c Falkmer et al., 2013
d Hirota et al., 2018
e Levy et al., 2020
f Soto et al., 2015
g Stuart & Lee 2017
h Sun et al., 2012
i Wang et al., 2020
j Wigham et al. 2019



*Sometimes categorised as screening and sometimes described as diagnostic

Key

Reporting of statistics

Review only reported interpretation of the statistic for the psychometric property

Review reported 'hard' statistic for the psychometric property

Number of primary studies

The number of primary studies included in the review paper that investigated that psychometric property

Methodological quality of the review papers

Review paper scored between 9-10 on JBI checklist

Review paper scored between 7-8 on JBI checklist

Review paper scored below 7 on JBI checklist

Methodological quality of the primary studies

High quality primary study(ies) with few concerns reported by review paper

Some of the primary studies were rated as high-quality and others were rated as low-quality by review paper

Low quality primary study(ies) with two or more concerns reported by review paper

Methodological quality of the primary study(ies) was not evaluated by review paper

Review papers

a Backes et al., 2014
b Bagdadli et al., 2017
c Falkmer et al., 2013
d Hirota et al., 2018
e Levy et al., 2020
f Soto et al., 2015
g Stuart & Lee 2017
h Sun et al., 2012
i Wang et al., 2020
j Wigham et al. 2019

Nine of the systematic review papers reported psychometric properties for screening measures (Backes et al., 2014; Hirota et al., 2018; Stuart & Lee 2017; Sun et al., 2012; Wang et al., 2020; Bagdadli et al., 2017; Soto et al., 2015; Wigham et al., 2019; Falkmer et al., 2013). Many of the screening measures scored satisfactory for sensitivity and specificity (ASDASQ; ASQ; ASRS; ASSQ; ATA; CAST; SCDC), however the evidence from the primary studies was either limited, unsatisfactory or not evaluated by the review papers. The ABC, AQ, SCQ, and SRS had psychometric properties reported and evaluated by a wider range of review papers but showed mixed results for sensitivity and specificity across different review papers. There were concerns highlighted over the sensitivity and reliability of the ABC, the specificity and construct validity of the AQ, and the specificity of the SCQ and SRS. Overall, there was very little high-quality evidence for the psychometric properties of the screening measures. There were no psychometric properties reported by the review papers for 11 of the screening measures: ACL, ASIEP, ASSERT, ASA, CSBQ, EQ, FBB-TES, KADI, MBAS, PDDRS and the WSQ.

Seven of the systematic review papers reported psychometric properties for diagnostic measures (Bagdadli et al., 2017; Falkmer et al., 2013; Wigham et al., 2019; Backes et al., 2014; Sun et al., 2012; Stuart & Lee 2017; Hirota et al., 2018). The ADOS-2, ADI-R, CARS and RAADS-R had psychometric properties reported and evaluated by a wider range of review papers. The ADOS-2 was rated as satisfactory across internal consistency, reliability and validity but showed mixed results in sensitivity and specificity. The ADI-R showed satisfactory results for internal consistency and validity, but there were mixed results in sensitivity, specificity, and reliability across two of the review papers (Sun et al., 2012; Wigham et al., 2019). The CARS scored satisfactory across internal consistency and

reliability. For the CARS there was mixed results for sensitivity and specificity across one of the review papers, but this evidence was not evaluated (Stuart & Lee 2017). The RAADS-R had mixed results for validity and specificity, but scored satisfactory in internal consistency, reliability, and sensitivity. There were no psychometric properties reported by the review papers for the ASDI and E2 -DC. Overall, there was limited high-quality evidence for the psychometric properties of the diagnostic measures.

Discussion/ Critique

Overlap and Datedness of the Literature

It is important for meta-reviews to evaluate the overlap and datedness of the literature (Hennessy et al., 2019). There was variation in the populations, contexts, and types of measures focussed on by the included reviews. Therefore, high overlap across primary research was not expected, due to the broad scope of the meta-review (Hennessy & Johnson, 2020). For two of the included systematic reviews (Bagdadli et al., 2017; Wigham et al., 2019), similar populations, contexts and types of measure were laid out in review aims, and very high overlap was found. Differences in overlap between these reviews was likely due to variation in the dates and operationalisation of concepts for inclusion and exclusion criteria (Hennessy & Johnson, 2020).

Half of the included review papers were over ten years old, which is older than the latest DSV-V and ICD-11 update in diagnostic criteria for ASD (Rosen et al., 2021; WHO 2018).Therefore the information and conclusions from these reviews must be interpreted with caution, as they may not be consistent with current diagnostic criteria. The results from the non-systematic review papers were far more dated than the included systematic reviews,

which may reflect the advances in systematic review methodology, updated guidelines and rigor towards conducting reviews over the last ten years (Page et al., 2021).

Quality of Review Papers

Critique of the Systematic Review Papers

Whilst the overall quality of the systematic review papers was reasonably high, there were some areas of consistent concern highlighted by the evaluation against the JBI checklist (Aromataris et al., 2015). There were only a few systematic reviews that addressed the likelihood of publication bias. This could have impacted the results of the reviews, as primary studies with promising or statistically significant positive results, are more likely to be published (Nair, 2019). Where the included reviews did not consider publication bias, the quality of the measurement and psychometric properties for ASD measures could appear to be stronger, and these results need to be interpreted cautiously (Ayorinde et al., 2020). Over half of the systematic reviews did not appear to use two independent reviewers for critical appraisal and data extraction, and therefore there was a higher potential for bias in the critical appraisal and data extraction (Drucker et al., 2016).

Over half of the systematic reviews did not have appropriate or standardised criteria for appraising the primary studies that were included. As such, the quality of the primary studies that form the evidence base within these reviews was not fully evaluated, which limits the conclusions that can be drawn (Drucker et al., 2016). The psychometric properties reported by these systematic reviews need to be interpreted cautiously as the methodological quality and risk of bias within the primary studies was unknown. Evaluating methodological quality and risk of bias using standardised tools is particularly important for validation and

diagnostic accuracy studies to provide confidence in the results and conclusions (Viswanathan et al., 2018).

Strengths of the systematic review evidence base included the appropriate use of the inclusion criteria in relation to review questions, and clear evidence of an appropriate and comprehensive search strategy with the use of relevant electronic databases (Aromataris et al., 2015). This meant that the systematic reviews were more likely to identify all the available evidence and that relevant studies were more likely to be included for their research questions (Cooper et al., 2018). This gives confidence in the identification of relevant ASD measures by the included review studies, and less chance that a measure for ASD could have been missed.

Critique of the Non-Systematic Review Papers

Although, the quality rating for the non-systematic review papers appeared to be high in response to the six broad categories of the SANRA scale, important limitations were identified. There were concerns about the lack of detail with regards to search strategy, search terms or inclusion criteria for most of the included non-systematic reviews. In many cases, there was limited information about how authors chose the studies that were included and none of these reviews had a list of included primary studies. Whilst a non-systematic review is not required to have as strict systematic criteria for the search strategy (Ferrari, 2015), having little information about how the literature was found, left the reviews open to bias, which could have influenced the results (Golder et al., 2008). This also made it more difficult to know if there were relevant studies that could have been missed. Insufficient detail and evaluation of primary studies meant that these reviews had to be excluded from the calculation of overlap and synthesis of psychometric properties, due a lack of clarity with

regard to where information was coming from (Hennessy et al., 2019; Hennessy & Johnson, 2020). In the absence of more formal guidelines, such as the PRISMA guidelines for systematic reviews (Page et al., 2021), there was also greater heterogeneity in the narrative methods used to synthesis data within the non-systematic reviews. Therefore, caution needs to be taken when considering the results from the non-systematic reviews.

Identified Autism Measures

There were many screening and diagnostic measures for ASD identified from the reviews. The existence of various measures is likely to reflect the complexity and variability in the presentation of ASD, which includes different levels of symptom severity and association with co-occurring conditions (Charman & Gotham, 2013). There were more screening than diagnostic measures identified for ASD, and these were generally shown to be quicker and easier to administrate (Sobieski et al., 2022). Screening measures provide a low-cost option for research and as an initial step in the identification and diagnosis of ASD in clinical settings (Marlow et al., 2019).

Characteristics provided by Review Papers

Most of the included review papers provided information about the characteristics of the identified ASD measures. For reviews that included both screening and diagnostic measures for ASD, there was also information about whether the measure was categorised as screening or diagnostic. Within the literature on autism there is a lack of a clear definition of what constitutes a diagnostic measure for ASD (Charman & Gotham, 2013). Clear guidelines for the diagnosis of ASD exist, but these are limited to specific certain countries, such as the National Institute for Health and Care Excellence (NICE) guidelines within the UK, which

recommends a small number of ASD diagnostic measures alongside multidisciplinary assessment (NICE, 2012). Within the review papers there was heterogeneity about which ASD measures were considered to be diagnostic. Six of the identified measures: the AAA, ASD-DC, ASDS, CARS, GARS and the RAADS-R were defined as screening measures in some review papers, and as diagnostic in other review papers. This highlights the importance of defining terms more clearly and consistently, and the need for guidelines that distinguish between diagnostic and screening measures (Charman & Gotham, 2013; Strunk et al., 2017).

Psychometric Properties

Within the literature on autism there have been many of individual validation studies that have evaluated the psychometric properties of the CARS, ADOS-2, and ADI-R (Volker & Lopata, 2008). The ADOS and ADI-R are often heralded as the gold standard for diagnostic assessment of ASD (Kamp-Becker et al., 2013). The results from the included systematic reviews showed that the psychometric properties of the ADOS-2 and ADI-R had been explored by a wider range of primary studies, but there were still mixed results about sensitivity and specificity, and concerns about the quality of the evidence base. This highlighted that even diagnostic tools considered ‘gold-standard’ show a large variation in diagnostic capabilities (Lefort-Besnard et al., 2020). In line with previous research, the ADOS-2 and the ADI-R were the most recommended diagnostic measures for the assessment of ASD based on the evidence for reliability and validity (Volker & Lopata, 2008). The CARS was identified as a useful measure, particularly by reviews conducted in LMICs, due to promising psychometric properties and ease of administration (Samms-Vaughan et al., 2017). However, the unknown methodological quality of the studies evaluating the CARS

limits the conclusion that can be drawn regarding this measure, and recent research has questioned the measurement invariance of the CARS (Stevanovic et al., 2021).

Psychometrics on reliability and validity outside of sensitivity and specificity were not widely reported by many of the review papers, apart from Bagdadli et al., (2017). There is a need for more thorough investigation of a wider range psychometric properties for ASD screening and assessment measures. The results for many of the psychometric properties were based on one or two primary studies. Considering that there were often concerns about the methodological quality of these primary studies, or that the primary studies were not critically appraised, this limits the conclusions that can be drawn. There is need for more thorough investigation of psychometric properties from high-quality primary studies for many of the ASD screening and diagnostic measures. Where there is limited information or concerns about the psychometric properties of measures, literature on autism has highlighted the need for triangulating information sources and the importance of multidisciplinary assessment alongside standardised diagnostic measures (Strunk et al., 2017; Wigham et al., 2018).

Cultural Adaptation, Translation and Validation

The reviews contained little information about the nations in which ASD measures were validated, and which languages they were available in. One review specifically focussed on the cultural adaptation of ASD measures, and therefore information about translation and cultural adaption was included (Soto et al., 2015). However, this was missing in many of the reviews. Information about translation and validation is important to help clinicians and researchers around the world select the most appropriate ASD measure for their geographical

location. This is especially important considering differences in language and culture have the potential to influence access to, interpretation and use of ASD measures (Matson et al., 2017). Four of the included reviews (Backes et al., 2014; Sun et al., 2013; Wang et al., 2020; Stewart & Lee, 2017) were focused on LMIC countries that are normally excluded by reviews focussed on Western developed countries. The reviews focussed on LMIC countries contained more information about where the included measures were validated, and which languages the measures had been translated into. There is a need for a more comprehensive review that identifies where different ASD measures have been translated and validated globally, and this presents an important opportunity for future research.

Information about translation and validation is further important to increase opportunities for cross-cultural research, interventions, and activities associated with autism. For example, this could help to increase opportunities for autistic individuals in international para-sport competitions, as there is a currently barrier of how to assess global eligibility for autistic athletes without a comorbid intellectual disability (International Paralympic Committee, 2016). As a fundamental aspect of eligibility in para-sport requires valid and reliable measures of impairment (Hutchinson et al., 2020), the results from this meta-review alongside further information on translation and validation could help to inform autism eligibility processes in a global sporting context. This presents another area for future research.

Limitations

There were several limitations to this meta-review. Firstly, the inclusion of articles only published in English and the use of Western literature databases, such as PsychInfo. This meant that the number of reviews conducted in non-western countries may have been

underestimated, and autism measures created and published in non-western regions of the world were less likely to be identified.

This meta-review focussed on ASD measures for use with older children, adolescents, and adults above the age of six years. Therefore, measures that were solely for use with under six years were excluded. This decision was based on differences in the types of ASD measures for young children which are more focussed on observational assessment. However, this could mean that potentially useful ASD measures may have been excluded.

Psychometric data was synthesised from the review papers, as it was not pragmatically possible to go into each of the individual primary studies to gather psychometric data or to assess the quality of the primary studies. Whilst harvest plots are a useful tool to visually present diverse findings (Crick et al., 2015), the use of harvest plots to synthesise psychometric data had several limitations. Harvest plots can be difficult to interpret (Burns et al., 2018), and did not include information about differences in the populations of the primary studies drawn on by the review papers. This limited what could be concluded about the psychometric properties and may have helped to explain differences in the findings of the psychometric data.

Hennessy et al. (2019)'s best practice guidelines for systematic meta-reviews were followed as far as was pragmatically possible. However, this meta-review was limited by having one reviewer for the literature search and data extraction. The use of two independent raters for each of these processes reduces the likelihood of bias and human error (Belur et al., 2021). In the absence of rigorous critical appraisal tools for non-systematic reviews, the SANRA scale was used to assess the quality of non-systematic reviews. The SANRA scale is more subjective, due to the broad questions that are more open to interpretation than the

detailed questions (and example ratings) found within the JBI checklist. Although the SANRA scale was used by the author on five practice reviews before assessing the reviews included in this study, it is likely that the SANRA quality ratings had a level of subjectivity.

Conclusion

This meta-review has identified ten systematic reviews and ten non-systematic reviews that evaluated screening and diagnostic measures for ASD in older children, adolescents, and adults. The quality of the systematic reviews was reasonably high, but there were concerns around approaches to critical appraisal, data extraction and publication bias. The quality of the non-systematic reviews was lower, with the main concerns for these reviews being the limited detail of the search strategy, a need for clearer and more explicit aims, and the datedness of the literature. There were 28 screening and 15 diagnostic measures for ASD that were identified by the included reviews for use with older children, adolescents, and adults. The characteristics and psychometric properties of the screening and diagnostic measures were narratively synthesised. In line with previous research, the ADOS-2 and ADI-R were the measures most recommended by the reviews, however there was variability in reports of validity and reliability. The lack of critical appraisal of primary studies by the included reviews meant that conclusions around psychometric properties had to be treated cautiously. There is a need for a more thorough investigation of a wider range psychometric properties for ASD screening and assessment measures. There was little information within the reviews about where ASD measures have been validated or what languages they have been translated into. This presents an important direction for future research.

References

Al Maskari, T. S., Melville, C. A., & Willis, D. S. (2018). Systematic review: cultural adaptation and feasibility of screening for autism in non-English speaking countries. *International journal of mental health systems*, 12(1), 1-19.

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(2), 202-212.

Alsayouf, H. A., Talo, H., Biddappa, M. L., Qasaymeh, M., Qasem, S., & De Los Reyes, E. (2020). Pharmacological intervention in children with autism spectrum disorder with standard supportive therapies significantly improves core signs and symptoms: a single-center, retrospective case series. *Neuropsychiatric Disease and Treatment*, *16*, 2779.

American Psychiatric Association (APA). (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5).* Washington, DC: American Psychiatric Publishing.

Aromataris E, Fernandez R, Godfrey C, Holly C, Kahlil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an Umbrella review approach. *Int J Evid Based Healthc.* 2015;13(3):132-40.

Backes, B., Mônego, B. G., Bosa, C. A., & Bandeira, D. R. (2014). Psychometric properties of assessment instruments for autism spectrum disorder: a systematic review of Brazilian studies. *Jornal Brasileiro de Psiquiatria*, *63*, 154-164.

Baer, L., & Blais, M. A. (2010). *Handbook of clinical rating scales and assessment in psychiatry and mental health* (pp. 7-35). New York: Humana Press.

Baethge, C., Goldbeck-Wood, S., & Mertens, S. (2019). SANRA—a scale for the quality assessment of narrative review articles. *Research integrity and peer review*, *4*(1), 1-7.

Baghdadli, A., Russet, F., & Mottron, L. (2017). Measurement properties of screening and diagnostic tools for autism spectrum adults of mean normal intelligence: A systematic review. *European Psychiatry*, *44*, 104-124.

Balasco, L., Provenzano, G., & Bozzi, Y. (2020). Sensory abnormalities in autism spectrum disorders: a focus on the tactile domain, from genetic mouse models to the clinic. *Frontiers in Psychiatry*, *10*, 1016.

Ballabriga MCJ, Escudé RMC, Llaberia ED. Escala d'avaluació dels trets autistes (A.T.A.): validez y fiabilidad de una escala para el examen de las conductas autistas. Rev Psiquiatr Infanto-Juv. 1994;4:254-63.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders*, *34*(2), 163-175.

Baron-Cohen, S., Wheelwright, S., Robinson, J., & Woodbury-Smith, M. (2005). The adult Asperger assessment (AAA): a diagnostic method. *Journal of autism and developmental disorders*, *35*(6), 807-819.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism,

malesand females, scientists and mathematicians. *Journal of autism and developmental disorders*, *31*(1), 5-17.

Belur, J., Tompson, L., Thornton, A., & Simon, M. (2021). Interrater reliability in systematic review methodology: exploring variation in coder decision-making. *Sociological methods & research*, *50*(2), 837-865.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: diagnostic validity. *The British Journal of Psychiatry*, *175*(5), 444-451.

Bishop, D. V. (1998). Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*(6), 879-891.

Bishop, S. L., Huerta, M., Gotham, K., Alexandra Havdahl, K., Pickles, A., Duncan, A., ... & Lord, C. (2017). The autism symptom interview, school-age: A brief telephone interview to identify autism spectrum disorders in 5-to-12-year-old children. *Autism Research*, *10*(1), 78-88.

Bordin, I. A., Rocha, M. M., Paula, C. S., Teixeira, M. C. T., Achenbach, T. M., Rescorla, L. A., & Silvares, E. F. (2013). Child Behavior Checklist (CBCL), Youth Self-Report (YSR) and Teacher's Report Form (TRF): an overview of the development of the original and Brazilian versions. *Cadernos de saúde pública*, *29*, 13-28.

Brugha, T. S., McManus, S., Smith, J., Scott, F. J., Meltzer, H., Purdon, S., ... & Bankart, J. (2012). Validating two survey methods for identifying cases of autism spectrum disorder among adults in the community. *Psychological medicine*, *42*(3), 647-656.

Campbell, J. M. (2005). Diagnostic assessment of Asperger's disorder: A review of five third-party rating scales. *Journal of Autism and Developmental Disorders*, *35*(1), 25-35.

Carpenter, P. (2012). Diagnosis and assessment in autism spectrum disorders. *Advances in Mental Health and Intellectual Disabilities*.

Castelloe, P., & Dawson, G. (1993). Subclassification of children with autism and pervasive developmental disorder: a questionnaire based on Wing's subgrouping scheme. *Journal of Autism and Developmental Disorders*, *23*(2), 229-241.

Charak, D. A., & Stella, J. L. (2002). Screening and diagnostic instruments for identification of autism spectrum disorders in children, adolescents, and young adults: A selective review. *Assessment for Effective Intervention*, *27*(1-2), 5-17.

Charman, T., & Gotham, K. (2013). Measurement Issues: Screening and diagnostic instruments for autism spectrum disorders–lessons from research and practise. Child and adolescent mental health, 18(1), 52-63.

Charman, T., Baird, G., Simonoff, E., Loucas, T., Chandler, S., Meldrum, D., & Pickles, A. (2007). Efficacy of three screening instruments in the identification of autistic-spectrum disorders. *The British Journal of Psychiatry*, *191*(6), 554-559.

Christensen, D. L., Braun, K. V. N., Baio, J., Bilder, D., Charles, J., Constantino, J. N., ... & Yeargin-Allsopp, M. (2018). Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2012. MMWR Surveillance Summaries, 65(13), 1.

Clancy, H., Dugdale, A., & Rendle-Short, J. (1969). The diagnosis of infantile autism. *Developmental Medicine and Child Neurology, 11*(4), 432–442.

Cohen, I. L., Schmidt-Lackner, S., Romanczyk, R., & Sudhalter, V. (2003). The PDD Behavior Inventory: a rating scale for assessing response to intervention in children with pervasive developmental disorder. *Journal of Autism and developmental Disorders, 33*(1), 31-45.

Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale: SRS-2* (p. 106). Torrance, CA: Western psychological services.

Cooper, C., Booth, A., Varley-Campbell, J., Britten, N., & Garside, R. (2018). Defining the process to literature searching in systematic reviews: a literature review of guidance and supporting studies. *BMC medical research methodology*, *18*(1), 1-14.

Dalton, J. E., Bolen, S. D., & Mascha, E. J. (2016). Publication bias: the elephant in the review. *Anesthesia and analgesia*, *123*(4), 812.

Derks, O., Heinrich, M., Brooks, W., Sterkenburg, P., McCarthy, J., Underwood, L., & Sappok, T. (2017). The Social Communication Questionnaire for adults with intellectual disability: SCQ-AID. *Autism Research*, *10*(9), 1481-1490.

Drucker, A. M., Fleming, P., & Chan, A. W. (2016). Research techniques made simple: assessing risk of bias in systematic reviews. *Journal of Investigative Dermatology*, *136*(11), e109-e114.

Duku, E., Vaillancourt, T., Szatmari, P., Georgiades, S., Zwaigenbaum, L., Smith, I. M., ... & Bennett, T. (2013). Investigating the measurement properties of the social

responsiveness scale in preschool children with autism spectrum disorders. *Journal of autism and developmental disorders*, *43*(4), 860-868.

Ehlers, S., Gillberg, C., & Wing, L. (1999). A screening questionnaire for Asperger syndrome and other high-functioning autism spectrum disorders in school age children. *Journal of autism and developmental disorders*, *29*(2), 129-141.

Elder, J. H., Kreider, C. M., Brasher, S. N., & Ansell, M. (2017). Clinical impact of early diagnosis of autism on the prognosis and parent–child relationships. *Psychology research and behavior management*.

Elwin, M., Schröder, A., Ek, L., & Kjellin, L. (2016). Development and pilot validation of a sensory reactivity scale for adults with high functioning autism spectrum conditions: Sensory Reactivity in Autism Spectrum (SR-AS). *Nordic Journal of Psychiatry*, *70*(2), 103-110.

Eriksson, J. M., Andersen, L. M., & Bejerot, S. (2013). RAADS-14 Screen: validity of a screening tool for autism spectrum disorder in an adult psychiatric population. *Molecular Autism*, *4*(1), 1-11.

Falkmer, T., Anderson, K., Falkmer, M., & Horlin, C. (2013). Diagnostic procedures in autism spectrum disorders: a systematic literature review. *European child & adolescent psychiatry*, *22*(6), 329-340.

Ferrari, R. (2015). Writing narrative style literature reviews. *Medical Writing*, *24*(4), 230-235.

Gabbay-Dizdar, N., Ilan, M., Meiri, G., Faroy, M., Michaelovski, A., Flusser, H., ... & Dinstein, I. (2022). Early diagnosis of autism in the community is associated with marked improvement in social symptoms within 1–2 years. *Autism*, *26*(6), 1353-1363.

Garnett, M. S., & Attwood, A. J. (1998). The Australian scale for Asperger's syndrome. *Asperger's syndrome. A guide for parents and professionals*, 17-19.

Gillberg, C., Gillberg, C., Råstam, M., & Wentz, E. (2001). The Asperger Syndrome (and high-functioning autism) Diagnostic Interview (ASDI): a preliminary study of a new structured clinical interview. *Autism*, *5*(1), 57-66.

Gilliam, J. E. (1995). *GARS: Gilliam autism rating scale*. Pro-ed.

Gilliam, J. E. (2001). *Gilliam Asperger's Disorder Scale: GADS*. Austin, TX: Pro-Ed.

Golder, S., Loke, Y., & McIntosh, H. M. (2008). Poor reporting and inadequate searches were apparent in systematic reviews of adverse effects. *Journal of Clinical Epidemiology*, *61*(5), 440-448.

Goldstein, S., & Naglieri, J. A. (2009). *Autism spectrum rating scales (ASRS)*. North Tonawanda, NY: Multi-Health System

Grodberg, D., Weinger, P. M., Halpern, D., Parides, M., Kolevzon, A., & Buxbaum, J. D. (2014). The autism mental status exam: sensitivity and specificity using DSM-5 criteria for autism spectrum disorder in verbally fluent adults. *Journal of autism and developmental disorders*, *44*(3), 609-614.

Hirota, T., So, R., Kim, Y. S., Leventhal, B., & Epstein, R. A. (2018). A systematic review of screening tools in non-young children and adults for autism spectrum disorder. *Research in Developmental Disabilities*, *80*, 1-12.

Hoekstra, R. A., Vinkhuyzen, A. A., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., ... & Van Der Sluis, S. (2011). The construction and validation of an abridged version of the autism-spectrum quotient (AQ-Short). *Journal of autism and developmental disorders*, *41*(5), 589-596.

Horwitz, E. H., Schoevers, R. A., Ketelaars, C. E. J., Kan, C. C., Van Lammeren, A. M. D. N., Meesters, Y., ... & Hartman, C. A. (2016). Clinical assessment of ASD in adults using self-and other-report: psychometric properties and validity of the Adult Social Behavior Questionnaire (ASBQ). *Research in Autism Spectrum Disorders*, *24*, 17-28.

Howlin, P. (2000). Assessment instruments for Asperger syndrome. *Child Psychology and Psychiatry Review*, *5*(3), 120-129.

Hull, L., Petrides, K. V., & Mandy, W. (2020). The female autism phenotype and camouflaging: A narrative review. Review Journal of Autism and Developmental Disorders, 7, 306-317.

Hutchinson, M. J., Phillips, J. L. K., Mason, B. S., Goosey-Tolfrey, V. L., & Beckman, E. M. (2020). Measures of impairment applicable to the classification of Paralympic athletes competing in wheelchair sports: A systematic review of validity, reliability and associations with performance. *Https://Doi.Org/10.1080/02640414.2020.1815957*, *39*(sup1), 40–61. https://doi.org/10.1080/02640414.2020.1815957

International Paralympic Committee. (2016). *International Standard for Eligible Impairments*. www.paralympic.org.

Kamp-Becker, I., Mattejat, F., Wolf-Ostermann, K., & Remschmidt, H. (2005). The Marburg Rating Scale for Asperger's Syndrome (MBAS)--a screening instrument for high-functioning autistic disorders. *Zeitschrift fur Kinder-und Jugendpsychiatrie und Psychotherapie*, *33*(1), 15-26.

Kanne, S. M., Christ, S. E., & Reiersen, A. M. (2009). Psychiatric symptoms and psychosocial difficulties in young adults with autistic traits. *Journal of autism and developmental disorders*, *39*(6), 827-833.

King, B. H., Navot, N., Bernier, R., & Webb, S. J. (2014). Update on diagnostic classification in autism. *Current opinion in psychiatry*, *27*(2), 105.

Klose, L. M., Plotts, C., Kozeneski, N., & Skinner-Foster, J. (2012). A review of assessment tools for diagnosis of autism spectrum disorders: Implications for school practice. *Assessment for Effective Intervention*, *37*(4), 236-242.

Krug, D. A., & Arick, J. R. (2003). *KADI. Krug Asperger's Disorders Index*. Pro-Ed.

Krug, D. A., Arick, J., & Almond, P. (1980). Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. *Child Psychology & Psychiatry & Allied Disciplines*.

Kurita, H., Koyama, T., & Osada, H. (2005). Autism-Spectrum Quotient–Japanese version and its short forms for screening normally intelligent persons with pervasive developmental disorders. *Psychiatry and clinical neurosciences*, *59*(4), 490-496.

Lau, W. Y. P., Kelly, A. B., & Peterson, C. C. (2013). Further evidence on the factorial structure of the Autism Spectrum Quotient (AQ) for adults with and without a clinical diagnosis of autism. *Journal of autism and developmental disorders*, *43*(12), 2807-2815.

Lauritsen, M. B. (2013). Autism spectrum disorders. *European child & adolescent psychiatry*, 22(1), 37-42.

Levy, S. E., Wolfe, A., Coury, D., Duby, J., Farmer, J., Schor, E., ... & Warren, Z. (2020). Screening tools for autism spectrum disorder in primary care: A systematic evidence review. *Paediatrics*, *145*(Supplement_1), S47-S59.

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. Journal of the American Academy of Child & Adolescent Psychiatry, 56(6), 466-474.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., ... & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of autism and developmental disorders*, *30*(3), 205-223.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of autism and developmental disorders*, *24*(5), 659-685.

Luteijn, E., Luteijn, F., Jackson, S., Volkmar, F., & Minderaa, R. (2000). The Children's Social Behavior Questionnaire for milder variants of PDD problems: Evaluation of

the psychometric characteristics. *Journal of autism and developmental disorders*, *30*(4), 317-330.

Marlow, M., Servili, C., & Tomlinson, M. (2019). A review of screening tools for the identification of autism spectrum disorders and developmental delay in infants and young children: recommendations for use in low-and middle-income countries. *Autism Research*, *12*(2), 176-199.

Matson, J. L., Hess, J. A., Mahan, S., & Fodstad, J. C. (2010). Convergent validity of the autism spectrum disorder-diagnostic for children (ASD-DC) and autism diagnostic interview-revised (ADI-R). *Research in Autism Spectrum Disorders*, *4*(4), 741-745.

Matson, J. L., Matheis, M., Burns, C. O., Esposito, G., Venuti, P., Pisula, E., ... & Goldin, R. (2017). Examining cross-cultural differences in autism spectrum disorder: A multinational comparison from Greece, Italy, Japan, Poland, and the United States. *European Psychiatry*, *42*, 70-76.

Matson, J. L., Nebel-Schwalm, M., & Matson, M. L. (2007). A review of methodological issues in the differential diagnosis of autism spectrum disorders in children. *Research in Autism Spectrum Disorders*, *1*(1), 38-54.

Matson, J. L., Terlonge, C., & Gonzalez, M. L. (2006). Autism Spectrum Disorders-Diagnosis (ASD-D). Baton Rouge, LA: Disability Consultants, LLC.

Matson, J. L., & Goldin, R. L. (2013). Comorbidity and autism: Trends, topics and future directions. *Research in Autism Spectrum Disorders*, *7*(10), 1228-1233.

Mayes, S. D., Calhoun, S. L., & Crites, D. L. (2001). Does DSM-IV Asperger's disorder exist?. *Journal of Abnormal child psychology*, *29*(3), 263-271.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ, 339*, 332-336.

Morgan, S. (1988). Diagnostic assessment of autism: A review of objective scales. *Journal of Psychoeducational Assessment, 6*(2), 139-151.

Myles, B. S., Bock, S. J., & Simpson, R. L. (2001). *ASDS: Asperger Syndrome Diagnostic Scale*. Pro-Ed.

National Collaborating Centre for Mental Health UK. (2012). *Autism: Recognition, referral, diagnosis and management of adults on the autism spectrum.* British Psychological Society.

Norris, M., & Lecavalier, L. (2010). Screening accuracy of level 2 autism spectrum disorder rating scales: A review of selected instruments. *Autism, 14*(4), 263-284.

Nylander, L., & Gillberg, C. (2001). Screening for autism spectrum disorders in adult psychiatric out-patients: a preliminary report. *Acta Psychiatrica Scandinavica*, *103*(6), 428-434.

Parks, S. L. (1983). The assessment of autistic children: A selective review of available instruments. *Journal of Autism and Developmental Disorders, 13*(3), 255-267.

Posserud, M. B., Breivik, K., Gillberg, C., & Lundervold, A. J. (2013). ASSERT–the autism symptom SElf-ReporT for adolescents and adults: bifactor analysis and validation in a large adolescent population. *Research in Developmental Disabilities, 34*(12), 4495-4503.

Reiersen, A. M., Constantino, J. N., Grimmer, M., Martin, N. G., & Todd, R. D. (2008). Evidence for shared genetic influences on self-reported ADHD and autistic symptoms in young adult Australian twins. *Twin Research and Human Genetics*, *11*(6), 579-585.

Rimland, B. (1971). The differentiation of childhood psychoses: An analysis of checklists for 2,218 psychotic children. *Journal of Autism & Childhood Schizophrenia*, *1*(2), 161–174.

Ritvo, R. A., Ritvo, E. R., Guthrie, D., Ritvo, M. J., Hufnagel, D. H., McMahon, W., ... & Eloff, J. (2011). The Ritvo Autism Asperger Diagnostic Scale-Revised (RAADS-R): a scale to assist the diagnosis of autism spectrum disorder in adults: an international validation study. *Journal of autism and developmental disorders*, *41*(8), 1076-1089.

Rosen, N. E., Lord, C., & Volkmar, F. R. (2021). The Diagnosis of Autism: From Kanner to DSM-III to DSM-5 and Beyond. *Journal of Autism and Developmental Disorders 2021 51:12*, *51*(12), 4253–4270. https://doi.org/10.1007/S10803-021-04904-1

Rutter, M., Bailey, A., & Lord, C. (2003). *The Social Communication Questionnaire.* Torrance, CA: Western Psychological Services.

Sappok, T., Heinrich, M., & Underwood, L. (2015). Screening tools for autism spectrum disorders. *Advances in Autism*, *1*(1), 12-29.

Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of autism and developmental disorders*, *10*(1), 91–103.

Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002). The CAST (Childhood Asperger Syndrome Test) Preliminary development of a UK screen for mainstream primary-school-age children. *Autism*, *6*(1), 9-31.

Sipes, M., & Matson, J. L. (2014). Factor structure for autism spectrum disorders with toddlers using DSM-IV and DSM-5 criteria. *Journal of autism and developmental disorders*, *44*(3), 636-647.

Skuse, D. H., Mandy, W. P., & Scourfield, J. (2005). Measuring autistic traits: heritability, reliability and validity of the Social and Communication Disorders Checklist. *The British Journal of Psychiatry*, *187*(6), 568-572.

Skuse, D., Warrington, R., Bishop, D., Chowdhury, U., Lau, J., Mandy, W., & Place, M. (2004). The developmental, dimensional and diagnostic interview (3di): a novel computerized assessment for autism spectrum disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *43*(5), 548-558.

Sobieski, M., Sobieska, A., Sekułowicz, M., & Bujnowska-Fedak, M. M. (2022). Tools for early screening of autism spectrum disorders in primary health care–a scoping review. *BMC primary care*, *23*(1), 1-26.

Song, Y., Fang, J., & Sun, L. (2009). A new, high precision behavior-based questionnaire for screening and pre-diagnosing autism in children: Implication for clinical applications and large population investigation. *Journal of Psychological Science*, 32(6), 1504–1507.

Soto, S., Linas, K., Jacobstein, D., Biel, M., Migdal, T., & Anthony, B. J. (2015). A review of cultural adaptations of screening tools for autism spectrum disorders. *Autism*, *19*(6), 646-661.

Stevanovic, D., Costanzo, F., Fucà, E., Valeri, G., Vicari, S., Robins, D. L., ... & Knez, R. (2021). Measurement invariance of the Childhood Autism Rating Scale (CARS) across six countries. *Autism Research*, *14*(12), 2544-2554.

Stewart, L. A., & Lee, L. C. (2017). Screening for autism spectrum disorder in low-and middle-income countries: A systematic review. *Autism*, *21*(5), 527-539.

Stoesz, B. M., Montgomery, J. M., Smart, S. L., & Hellsten, L. A. M. (2011). Review of five instruments for the assessment of Asperger's disorder in adults. *The Clinical Neuropsychologist*, *25*(3), 376-401.

Sun, X., Allison, C., Auyeung, B., Matthews, F. E., Baron-Cohen, S., & Brayne, C. (2013). What is available for case identification in autism research in mainland China?. *Research in Autism Spectrum Disorders*, *7*(5), 579-590.

Thabtah, F., & Peebles, D. (2019). Early autism screening: a comprehensive review. *International journal of environmental research and public health*, *16*(18), 3502.

Volker, M. A., & Lopata, C. (2008). Autism: A review of biological bases, assessment, and intervention. *School Psychology Quarterly*, *23*(2), 258.

Wallis, K., & Pinto-Martin, J. (2008). The challenge of screening for autism spectrum disorder in a culturally diverse society. *Acta Paediatrica*, *97*(5), 539-540.

Wang, J., Hedley, D., Bury, S. M., & Barbaro, J. (2020). A systematic review of screening tools for the detection of autism spectrum disorder in mainland China and surrounding regions. *Autism*, *24*(2), 285-296.

Wigham, S., Rodgers, J., Berney, T., Le Couteur, A., Ingham, B., & Parr, J. R. (2019). Psychometric properties of questionnaires and diagnostic measures for autism spectrum disorders in adults: A systematic review. *Autism*, *23*(2), 287-305.

Wing, L., Leekam, S. R., Libby, S. J., Gould, J., & Larcombe, M. (2002). The diagnostic interview for social and communication disorders: Background, inter-rater reliability and clinical use. *Journal of child psychology and psychiatry*, *43*(3), 307-325.

World Health Organization. (2013). *Autism spectrum disorders and other developmental disorders: From raising awareness to building capacity.* Geneva, Switzerland: WHO Document Production Services.

World Health Organization. (2018). *The International Classification of Diseases, 11th edition (ICD-11).* Geneva, Switzerland: American College of Physicians.

Zahorodny, W., Shenouda, J., Mehta, U., Yee, E., Garcia, P., Rajan, M., & Goldfarb, M. (2018). Preliminary evaluation of a brief autism screener for young children. *Journal of Developmental and Behavioral Pediatrics, 39*(3), 183.

Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., ... & Elsabbagh, M. (2022). Global prevalence of autism: a systematic review update. *Autism Research*, *15*(5), 778-790.

ANNA EAST BA (Hons) MSc

SECTION B:

Increasing access to competitive sport: A Delphi study exploring the views of an expert panel on a method for establishing global eligibility criteria for a diagnosis of autism in a sporting context

Word Count: 7976 (481)

For the Journal of Sports Sciences

A thesis submitted in partial fulfilment of the requirements of Canterbury Christ Church University for the degree of Doctor of Clinical Psychology

APRIL 2023

SALOMONS INSTITUTE FOR APPLIED PSYCHOLOGY

CANTERBURY CHRIST CHURCH UNIVERSITY

Abstract

There are currently limited opportunities for autistic individuals in para-sport, as competitions for autistic athletes do not exist unless athletes have a comorbid intellectual disability. Participation in para-sport requires proof of eligible impairment as a first step towards a global eligibility process to facilitate access to international competition. This study aimed to examine the views of an expert panel on a method for establishing global eligibility criteria for a diagnosis of autism in a sporting context. Twenty-seven international participants took part in a three-round Delphi panel using online surveys. The results of the study showed that there was high consensus around a gold standard process of eligibility, which included an agreed definition of impairment and evidence that would need to be provided. There were lower levels of consensus and agreements around whether there should be an alternative process for countries that are unable to access the gold standard. Key challenges and barriers were identified including social and cultural differences, attention to co-morbidity and the heterogeneity of autism. Further research is needed to explore how autism impacts performance during competition in particular sports, and to develop classification and minimum impact criteria to facilitate access to international competition.

Keywords: Autism; Global; Eligibility; Para-sport; Delphi.

Introduction

Autism spectrum disorder (ASD), or autism spectrum condition (ASC), is a neurodevelopmental condition characterised by impairments in social communication and interaction, restricted or repetitive patterns of behaviour or interests, and sensory difficulties (American Psychiatric Association [APA], 2013). ASD is used throughout this study to ensure consistency, as this is the most common term for autism globally and the term currently used by the World Health Organisation (WHO, 2023). Many autistic individuals[1] are at greater risk of co-occurring physical and mental health conditions (Rydzewska et al., 2021; Sala et al., 2020; Weir et al., 2021; Hossain et al., 2020). A diagnosis of ASD can mean that individuals have additional challenges with everyday activities (Operto et al., 2021; Saulnier & Klaiman, 2022) and motor skills deficits (Dziuk et al., 2007; Mohd Nordin et al., 2021).

Autism and Physical Activity

The higher prevalence of difficulties and co-occurring conditions has been shown to significantly impact the quality of life for individuals with ASD (Scharoun et al., 2017). Physical activity (PA) participation is one modifiable risk factor that can affect health outcomes and quality of life (Healy et al., 2018). There are numerous benefits of PA and sport for autistic individuals (Colombo-Dougovito et al., 2020; Cunningham, 2019; Tamminen et al., 2020; Tarr et al., 2020). PA is associated with improvements in sleep (Tse et al., 2019), social communication and interactions (Colombo-Dougovito et al., 2020), cognition (Tan et al., 2016), motor control (Ketcheson et al., 2017; Pan et al., 2017), and physical health and fitness (Pan, 2011). Whilst benefits have been researched in the short-term, there is a paucity of research evidence about the long-term effects of PA-based interventions (Sefen et al., 2020).

[1] 'Identity first' (autistic people) rather than the 'person first' (people with autism) language is used in response to the strong preferences of autistic people advocating for this terminology (https://autisticadvocacy.org/about-asan/identity-first-language/; (Vivanti, 2020).

Despite the benefits of PA, research has shown that autistic individuals are less active and more likely to lead sedentary lives, when compared to the general population (Jones et al., 2017; Stanish et al., 2017), and those with other disabilities (Case et al., 2020). Individuals with ASD may be less motivated and have fewer opportunities to participate in PA (Scharoun et al., 2017). As a result of decreased activity levels, individuals with ASD are more likely to be overweight than their typically developing counterparts (Thom et al., 2022; Zheng et al., 2017), thus leading to further health-related challenges (Scharoun et al., 2017).

Barriers to Physical Activity

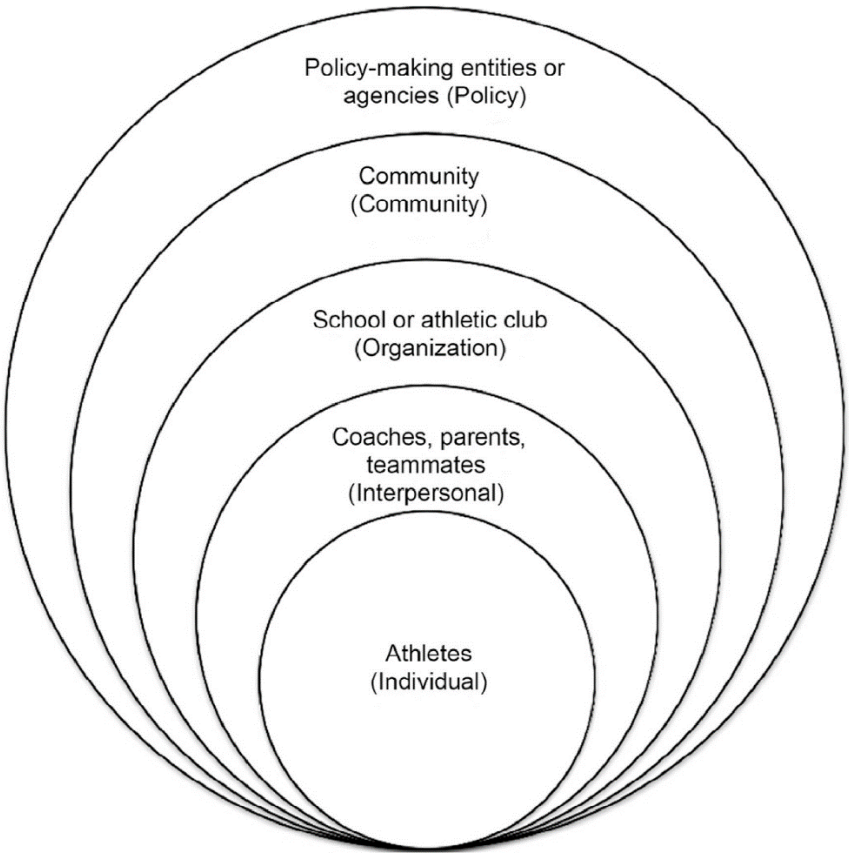
Several barriers have been identified that can impact the participation of autistic individuals in PA (Nichols et al., 2019). Finances, time, motor skills, behavioural and learning difficulties, need for supervision, lack of partner or teammates, lack of available transport, and lack of resources have been highlighted as potential barriers, alongside a lack of opportunities and access to PA and sport (Scharoun et al., 2017). However, research on PA opportunities, interventions, and benefits for autistic individuals has predominantly focussed on children and young people (Huang et al., 2020). Must et al. (2015) reported a positive relationship between age and the total number of barriers to PA and therefore, highlighted a need for more research to be conducted with adults.

Obrusnikova and Miccinello (2012) use the socioecological model adapted from Mcleroy et al., (1988), to categorize barriers to PA for autistic individuals. In this model, barriers are grouped into five categories: intrapersonal, interpersonal, physical, community and policy/ institutional (Figure 1). Intrapersonal barriers are the most frequently cited barriers to PA among autistic individuals (Must et al., 2015). However, all five of these categories need to be addressed to overcome the barriers to PA, including at the policy and institutional level (Obrusnikova & Miccinello, 2012). Little is known about how wider social,

systemic, and policy forces shape PA opportunities and access for autistic individuals, and therefore more research is needed in this area (Jachyra et al., 2020).

Figure 1

Adapted socioecological model from Bogardus et al. (2019)[2] *and applied to physical activity barriers*



Increasing Physical Activity through Sport

There have been calls to increase opportunities for PA globally, particularly for people with long term conditions (Bull et al., 2020). In 2018, the WHO recommended the

[2] Figure 2 taken from Bogardus et al. (2019) https://doi.org/10.1016/j.jshs.2017.11.001

creation of appropriate and supportive environments for physical activity for all population groups (WHO, 2018), and in 2020 WHO explicitly recommend PA as beneficial for people living with disability (Bull et al., 2020). In a sample of 35 autistic adolescents, Stanish et al., (2015) reported that half of the participants would like to do more PA than they were presently doing. Given the limited access to PA opportunities (Pan & Frey, 2006), efforts are needed to increase opportunities (Sefen et al., 2020). There have been calls to increase the number of high-quality PA programs tailored specifically to autistic individuals (Craig, 2021), to reduce barriers and facilitate inclusion (Bantjes & Swartz, 2018).

Organised sporting competitions are an important form of PA, and competition is a powerful social factor motivating people to increase activity (Vallerand, 2012). Although there is an abundance of evidence about health benefits of PA, there is a paucity of research about the opportunity to practice sport for autistic individuals (Vetri & Roccella, 2020). Vetri and Roccella (2020) highlighted two main barriers to practicing a sport for autistic individuals. The first was difficulties in social interaction and sensory stimulation particularly important in team sports (Smirni et al., 2019). The second was limitation in motor functions, as research has shown that autistic individuals have more difficulties in balance, gait, movement speed, motor control and joint flexibility when compared to individuals with typical development (Manicolo et al., 2019; Ozonoff et al., 2008). However, more research is needed in this area.

Para-sport Opportunities

Para-sport competitions are restricted entry competitions for people with an eligible impairment, promoting social inclusion and enabling fairer sporting competition (Bantjes & Swartz, 2018). Participation in para-sport can improve health and foster psychological well-being by providing opportunities for connection, and a sense of meaning and purpose in life (Vanderstraeten & Oomen, 2010). Although, many autistic individuals have competed

successfully in elite mainstream sport up to Olympic level, research has shown that a para-sport context can be beneficial for some autistic individuals (Duquesne et al., 2022). A para-sport context could reduce potential barriers and obstacles to mainstream sport (Duquesne et al., 2022). However, there are currently limited opportunities for autistic individuals in para-sport, as competitions for autistic athletes do not exist unless athletes have a comorbid intellectual disability (International Paralympic Committee, 2016). There is currently a barrier of how to assess eligibility for autistic athletes without a comorbid intellectual disability (International Paralympic Committee, 2016). When considering the socioecological model, this presents an institutional barrier to participation (Obrusnikova & Miccinello, 2012), and there is a need to increase opportunities for autistic athletes to participate in organised para-sport competition.

Evidence Based Eligibility and Autism Assessments

Participation in para-sport requires proof of eligible impairment (Tweedy & Vanlandewijck, 2011). Therefore, to enable participation, there needs to be an established method to determine eligibility, as a starting point to widen para-sport opportunities for autistic athletes. This involves working towards an internationally agreed process to confirm that an athlete has a diagnosis of autism (Tweedy & Vanlandewijck, 2011).

There are many different ways of assessing autism globally (Lord & Jones, 2012). The heterogeneity in the presentation of autism and the severity of symptomology (Rosen et al., 2021), as well as cultural and social differences in conceptualising ASD (Matson et al., 2017), has contributed to the development of a wide number of screening and diagnostic measures for ASD. These measures have been reviewed in Section A. The current diagnostic gold standard includes a multi-disciplinary assessment using the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview Revised (ADI-R) (Kamp-Becker et al., 2021). As demonstrated by the meta-review in Section A, this gold

standard of assessment is not available in every country due to access, resource, training, language, cultural adaptation and validation barriers (Samms-Vaughan et al., 2017). Assessments for autism are predominantly produced, researched and used in regional contexts (Hahler & Elsabbagh, 2015), which presents a challenge to determining eligibility in a global sporting context.

Taxonomy and Classification

This project drew on theories of taxonomy and evidence-based classification in autism, to understand the key difficulties and barriers in establishing a robust method of confirming eligibility in a global sporting context (King et al., 2014; Tweedy & Vanlandewijck, 2011). Taxonomic theory, including impact on functioning, is used as the basis of Paralympic eligibility classification systems (Tweedy & Vanlandewijck, 2011). This project was informed by previous studies researching eligibility and classification in para-sport (e.g. Runswick et al., 2021). Principles of eligibility were guided by previous research and a conceptual model presented by Van Biesen et al. (2021) of eligibility and classification for athletes with intellectual disabilities. Methodological considerations were informed by a study by Runswick et al., (2021), who utilised a Delphi panel to establish classification for footballers with visual impairments.

This project was conducted in partnership with the organisation Virtus. Virtus is the International Sports Federation for athletes with intellectual impairments, and operates an existing eligibility framework for athletes with intellectual disabilities (Van Biesen et al., 2021). Partnership with Virtus provided a para-sport context in which to begin the process of researching, testing and eventually implementing an eligibility process for autistic athletes.

Rationale and Study Aims

Research, into a method of confirming an autism diagnosis globally represents an important first step towards creating para-sporting competitions for autistic athletes and reducing barriers to participation. This project aimed to examine a method for establishing global eligibility criteria for a diagnosis of autism in a sporting context. The project sought to establish whether a panel of experts can reach a consensus opinion on the following research questions:

1. What is the gold standard process for confirming autism diagnosis in a global eligibility context for sport?
2. What are the dimensions that are important to consider when assessing autism in a global eligibility context for sport?
3. What are the key challenges or barriers that need to be overcome?
4. What method should be used as an alternative for nations that cannot comply with the gold standard?

Summary and Relevance to Psychology

This project aimed to widen access to competitive sport and increase PA opportunities for autistic individuals. Creating opportunities at the elite level of sport has the potential to diffuse through and increase opportunities at the grass-root level of sport (Craig, 2021). This aligns with National Health System (NHS) values of ‘everyone counts’ and ‘improving lives’, as increasing participation to para-sport for autistic individuals has the potential to affect health outcomes and challenge stigma (Bantjes & Swartz, 2018; Healy et al., 2018). Involvement in sport is a driver for social inclusion (Bantjes & Swartz, 2018), which has been recognised by the British Psychological Society (BPS) (2008) as an important psychological aim.

Method

Design

A three round Delphi (Dalkey & Helmer, 1963) was employed to explore consensus about establishing global eligibility criteria for a diagnosis of autism in a sporting context. The Delphi method consists of rounds of data collection from participants with 'expert' or lived experience in an area (Hasson et al., 2000). Widely used in health research to explore areas of limited or disputed research (Iqbal & Pipon-Young, 2009), the Delphi method has informed policy and planning (Jorm, 2015). This method shares feedback among participants to encourage consensus-building (Hsu & Sandford, 2007), a key reason this was chosen over other qualitative methods such as grounded theory. The Delphi method has been used in previous research on para-sport eligibility and classification (Runswick et al., 2021) and on physical activity benefits for people with long-term conditions (Reid et al., 2022).

This study utilised both qualitative and quantitative methods (Hsu & Sandford, 2007). To access an international sample, online surveys were used. Questions moved progressively from more open-ended questions in pilot-round (PR) and round-one (R1), to closed questions and Likert rating-scales in round-two (R2). In round-three (R3), participants were sent an individualised survey, displaying their R2 responses alongside the groups' responses. This allowed participants to review R2 ratings alongside the groups' responses, to refine 'expert' opinions (Hasson et al., 2000).

Recruitment

For this project, the term 'expert' was defined as "any individual with relevant knowledge and experience of a particular topic" (Cantrill et al., 1996, pg. 69). Participant inclusion and exclusion criteria are detailed in Table 1. Experts with a range of different experience including psychologists, medical professionals, academic researchers, sports eligibility officers, coaches, and parents of autistic athletes were recruited. Participants

needed to have at least three years of experience supporting autistic individuals, as well as meet one other inclusion criteria detailed in Table 1. It was important to include those with experience assessing autism, as well as those with experience of sports eligibility processes. Experts-by-experience, such as parents of autistic athletes, were recruited to provide first-hand recipient experiences on autism assessment and eligibility processes (Hardy et al., 2004).

Table 1

Participant inclusion and exclusion criteria

Participants need to meet criteria 1 and one of the other criteria
1. ***At least three years of experience supporting autistic people***
2. Involvement and experience using tests to assess Autism (e.g., as a psychologist who administers autism assessments)
3. Involvement and experience of sports eligibility and classification (e.g., as an eligibility officer for the Paralympics)
4. At least two years' experience coaching athletes with Autism
5. Experience of being an athlete with Autism
6. Experience of being a parent of an athlete with Autism
7. Access to professional networks related to Autism
Participants will be excluded if they:
Do not have sufficient English language skills to read and respond to a survey
Are unable to access a survey on an online platform
Are under 18 years old

Purposeful and snowball sampling strategies were used to ensure a range of different nationalities and experience. Potential participants were identified through key organisations (e.g., Virtus and the National Autistic Society), relevant published research, and contacts of the research supervisors. Potential participants were contacted by email, sent an invite letter (Appendix H), and asked to confirm which of the inclusion criteria they met.

Ethics

Full approval was received from Independent Research Review Panel and the Ethics Panel at Canterbury Christ Church University in October 2021 (Appendices I-J). This project followed the BPS Code of Ethics (2009). Participants were emailed an information sheet, describing the purpose, aims, benefits, and risks of the research, alongside participant data, confidentiality, and the withdrawal process (Appendix K). Informed consent was obtained (Appendix L), and a debrief sheet was emailed following R3 (Appendix M). Participants were informed that aggregated data and quotes would be shared anonymously with other participants, this report, and in future publications. Participant numbers were used to ensure anonymity, with participant names and contact details stored on a password protected database. Study data was stored on a separate password protected database.

Participants

There were 27 participants included in the study. There was an 89% completion rate across participants in PR, R1 and R2, and a 93% completion rate in R3. This meant that the final sample size was 25. One participant withdrew after R1, and the data previously collected remained in the study. Completion rates in each round were considered excellent, and higher than other Delphi studies (40-75%; Gordon, 1994). Table 2 details participant demographics including: age; gender; nationality; and job role. Figures 2 and 3 illustrate the experience of the panel relating to autism assessments, sport, coaching and eligibility processes. A total of 18 different countries were represented by the panel (Figure 4).

Data Collection and Analysis

The three-round Delphi process with a pilot-round, took 11 months between February 2022 and January 2023. The data collection and analysis for each round is described below. Figure 5 depicts the Delphi procedure flowchart.

Table 2

Participant demographics

Demographics			**N**	**Mean**
Age		18-24	-	44.7 years
		25-34	4	
		35-44	9	
		45-54	7	
		55-64	5	
		Did Not Answer	2	
Gender		Female	18	
		Male	9	
Nationality	Europe	British	4	
		British/ Polish	1	
		Belgium	1	
		Spanish/ Turkish	1	
		Irish	2	
		Russian	2	
		Polish	1	
		Portuguese	1	
		Icelandic	1	
	Americas	American	2	
		Canadian	1	
		Canadian/British	1	
		Brazilian	1	
		Mexican	1	
	Africa	South African	1	
	Asian	Japanese	1	
		Indian	3	
	Oceania	Australian	2	
Current Occupation		Academic, Professor or Researcher	8	
		Psychologist	9	
		Medical Doctor	1	

	Sports Coach or Director	2
	Sports Eligibility	1
	Disability Advisor	1
	Legal Advisor	1
	Full-time Parent	1
	Retired	1
	Did Not Answer	2

Figure 2

Experience of the panel: autism assessments

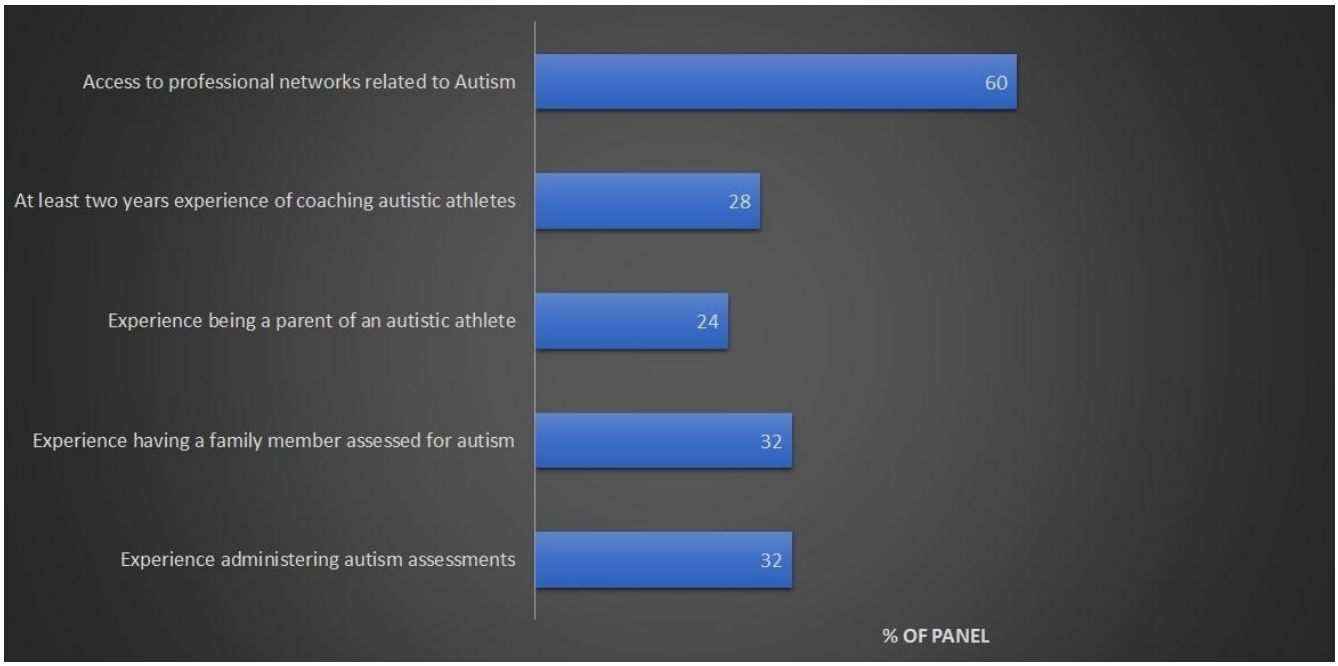


Figure 3

Experience of the panel: sport, coaching and eligibility processes

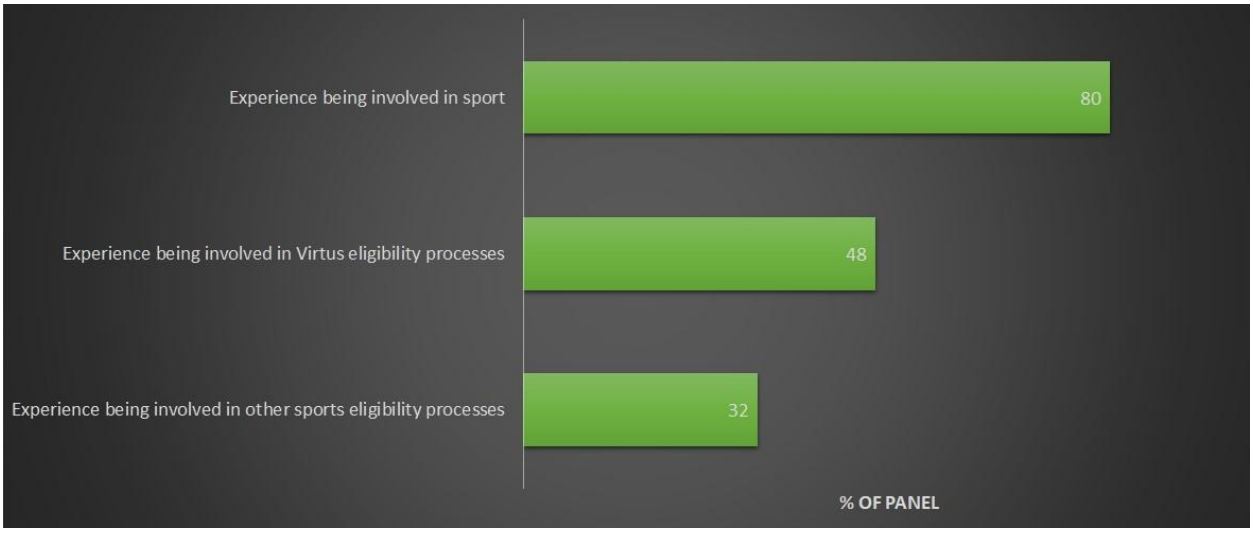
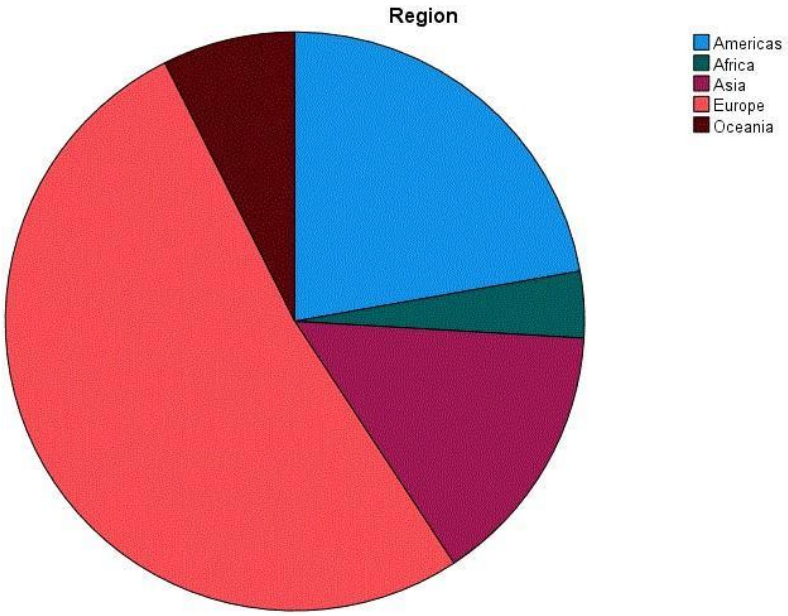


Figure 4

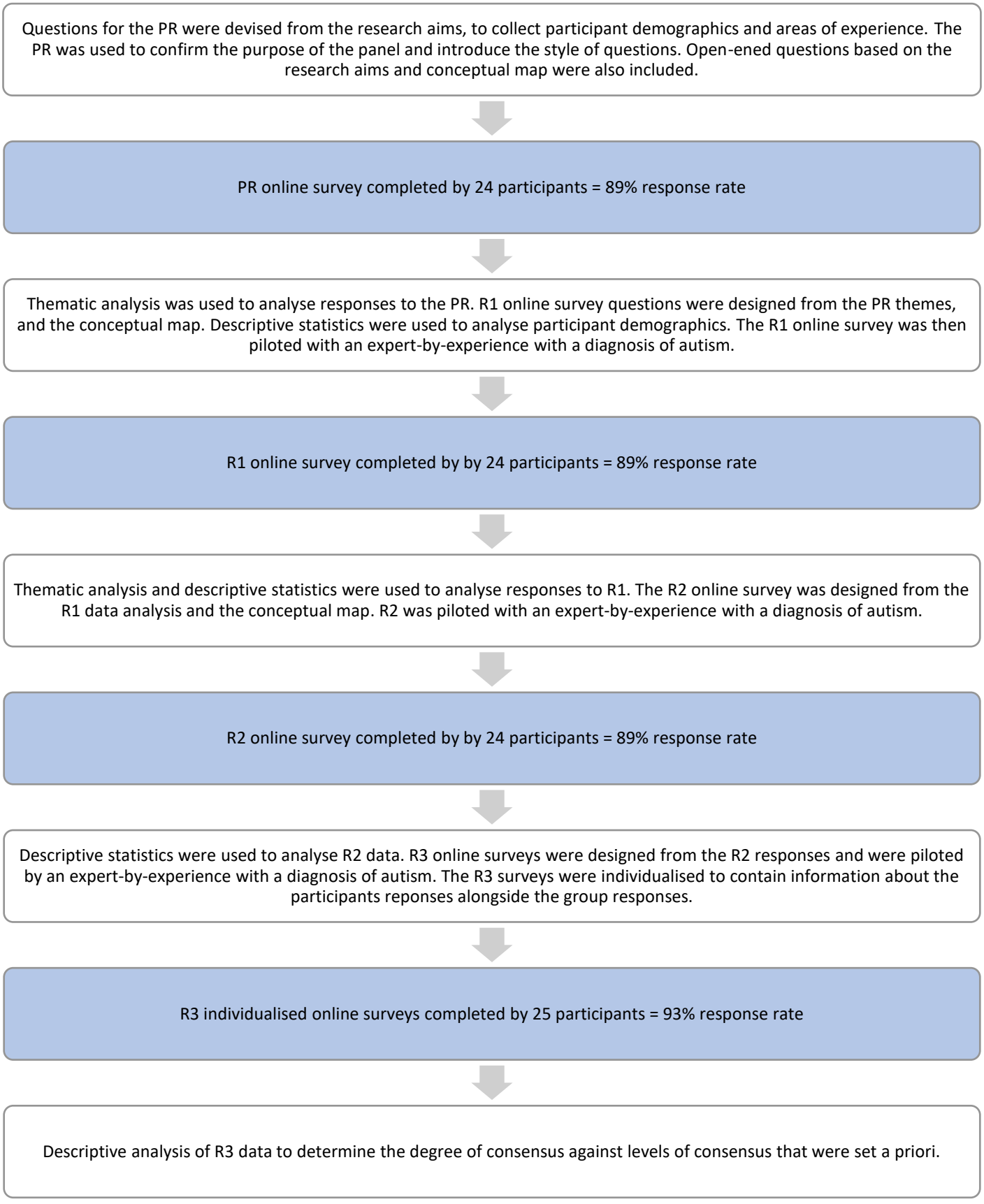
World regions represented by the panel



The researcher developed four online questionnaires to meet the research aims, distributed using Qualtrics^{XM} Online Survey software. As eligibility processes within para-sport have specific structures and regulations, the Delphi method was adapted to fit this topic of eligibility as seen in previous Delphi research on para-sport eligibility and classification (Runswick et al., 2021). Adaptations included giving participants detailed explanations with reference to current procedures, policy, or terminology that may not be familiar (Runswick et al., 2021). This was important due to the variety in participants with differing areas of expertise. Different response scales such as dichotomous and multiple-choice questions were used for areas and ideas that were introduced in R2, instead of the traditional Likert scale. This gave the panel opportunities to respond to the information given and influence the wording of the questions in R3. In R3, the 11-point rating scale was not always on an agree to disagree scale, but appropriate corresponding labels were used for each question to ensure that the responses fit with the wording of the question and broader eligibility terminology.

Figure 5

The Delphi procedure flowchart



A conceptual map (Appendix N) detailing the components of eligibility needing to be addressed, was produced to guide the development of the online surveys. This was produced in consultation with the research supervisor, who had extensive experience in para-sport eligibility processes, and was based on the eligibility process used for athletes with intellectual disabilities (Van Biesen et al., 2021).

PR Online Survey

Questions designed to measure participant demographics and experiences of autism assessments, eligibility processes, and sport were administered in the PR. The PR included three open-ended questions, informed by the research aims, about the most important factors, benefits, and difficulties or issues of a having a global eligibility process for autism (Appendix O). Thematic analysis (Braun & Clarke, 2006) of open-ended questions was conducted to identify overarching themes, in line with the aims of the research. Braun and Clark's (2006) iterative six step approach was followed; familiarization of data, coding, searching for themes, reviewing and checking themes by revisiting the coded extracts and full data set several times, defining and naming themes, and reporting the findings. PR responses were analysed inductively and deductively identifying 'theory-driven' and 'data-driven' codes (Appendix P; Booth & Carroll, 2015).

R1 Online Survey

R1 presented the aggregated demographics of the expert panel and the themes from PR responses. Questions were asked about the following areas: definitions of autism, eligibility processes and autism assessments (Appendix Q). These questions were informed by the PR themes and the conceptual map. There was a mixture of open-ended questions, ranking questions, and Likert rating-scale questions. R1 responses were analysed using thematic analysis and descriptive statistics.

R2 Online Survey

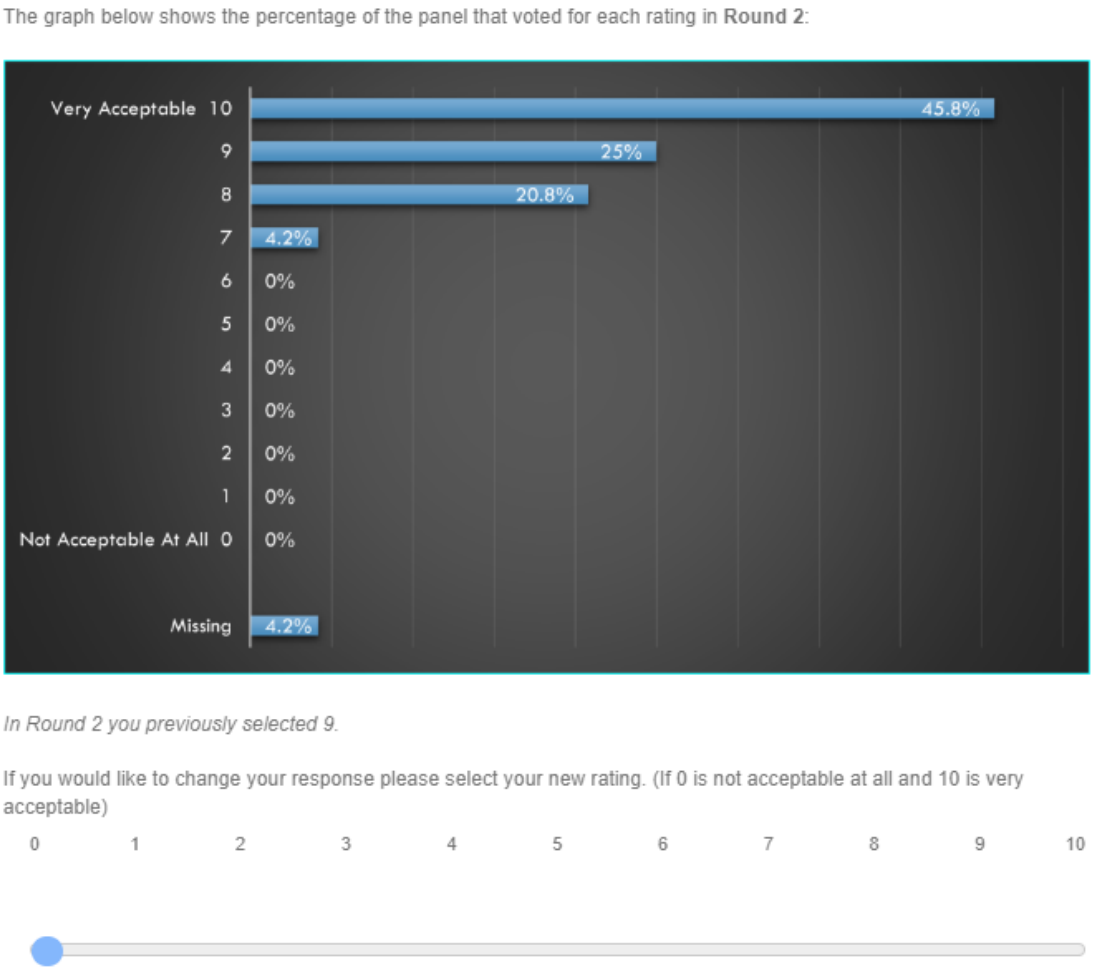
R2 presented the themes and descriptive statistics from questions in R1. Questions in R2 were informed by the responses in R1. A brief literature search was conducted to find the number of citations of the main validation paper for five autism diagnostic assessments, and the number of countries these assessments had some validation in (Appendices R-S). Closed questions in R2 asked about: definition of autism, what evidence is needed, autism assessments, and exceptions for countries that cannot meet the standards (Appendix T). Dichotomous and multiple-choice questions were used for areas and ideas that were introduced in this round. 11-point Likert rating-scale questions (from 0-10) were used for areas introduced in previous rounds, enabling the observation of smaller changes between rounds (Sharkey & Sharples, 2001). Follow up questions enabled participants to give qualitative comments or reasons for their responses.

R3 Online Survey

The purpose of the final round was to feed-back responses to R2 and invite participants to review their responses in the light of the overall group feedback. Qualitative comments made by participants in response to R2 were provided for clarification. R3 consisted of questions in the same areas as R2, but these were adapted and informed by responses in R2 (Appendix U). All questions in R3 were asked on an 11-point Likert rating-scale. Each participant was sent a personalised survey with R3 questions, the aggregated group responses for R2 questions, and their individual score for each R2 question (Figure 6).

Figure 6

Example of individualised R3 question

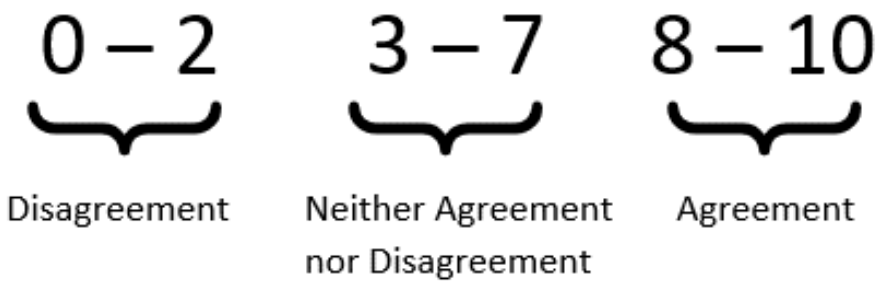


The Statistical Package for Social Science (SPSS, 2007) was used to carry out the descriptive analyses of the R3 data to determine levels of consensus. In line with other Delphi studies (Hackett et al., 2006; Vosmer et al., 2009), the overall percentages of disagreement and agreement were calculated to determine the strength of consensus in relation to each item. The 11-point rating scale was not always on an agree to disagree scale, but appropriate corresponding labels were used for each question. This included scales of not acceptable at all to very acceptable, not important at all to very important, and not desirable at all to very

desirable. Scores from each of these 11-point Likert scales were then collapsed into three bands of agreement to disagreement shown in Figure 7.

Figure 7

Collapsed categories from Likert scales



To address the problem of setting consensus too low, consensus was defined at three levels: high, medium, and low (Table 3). Levels of consensus were agreed a priori in line with Vosmer et al., (2009).

Table 3

Consensus levels

Consensus levels	**Percentage of agreement/ disagreement among participants in bands 0-2 and 8-10**
High consensus	≥80%
Medium consensus	65-79%
Low consensus	51-64%

In addition, medians and interquartile ranges (IQR) were calculated for each item (Jones & Hunter, 1995; Vosmer et al., 2009). Medians and IQR are affected less by extreme data values than means and standard deviations, offering offer valuable information on the

distribution and variability of the data (Marsh, 1998). For R3 analysis, participant data was divided into two groups. Group 1 (N=9) included participants with experience administering autism assessments, and Group 2 (N=16) included participants without experience administering autism assessments. This allowed for between-groups comparison, as well as comparison with overall degrees of consensus.

Quality Assurance and Reflexivity

Using a critical realist epistemological position (Bhaskar, 1979), acknowledged the existence of real-world structures, and that knowledge is mediated through human discourses and assumptions (Sayer, 2004). The use of qualitative and quantitative data from diverse participants provided a deeper understanding by allowing participants to expand on the reasons for their choices (Almalki, 2016; Patton, 1999). Online surveys facilitated efficient and anonymous data collection. Time between rounds helped with nuanced analysis (Braun & Clarke, 2006). Anonymity among participants aided balanced consideration of responses, reducing potential conformity biases (Bowles 1999; de Meyrick, 2003). Researcher subjectivity and preconceptions were considered through regular reflection with supervisors and the use of a research diary, to improve the decision-making throughout data collection and analysis and provide different perspectives on the process of research (Borg, 2016; Appendix V). The wording of questions in each round survey round was considered by three researchers to manage individual researcher subjectivity through discussion of different perspectives (Beiderbeck et al., 2021).

Service User Input

This project emerged from the request of autistic athletes and their families in Virtus, because of a lack of a specific para-sporting competition for those with autism. Preliminary work was piloted with experts and parents of autistic athletes, by field testing a Delphi style questionnaire. The surveys for each of the Delphi rounds were piloted by an

expert-by-experience with a diagnosis of autism, who also read a draft of this research report. Changes to simplify language and define terms were discussed with the research lead.

Results

Pilot and Round 1

Four main themes were identified, related to the important dimensions, processes, and challenges of confirming an autism diagnosis in a global eligibility context for sport. The first was the need for a **high-quality process** of eligibility, which considered qualities such as having a fair process of inclusion and exclusion, a clear and simple process, and international agreement on the process. The second theme highlighted the importance of using **standardised criteria and evidence,** which included having agreement on standardised assessments, formal diagnostic criteria and evidencing the impact on sports performance. The third theme identified was **accessibility,** which described increasing access to competitive sport, and the need for assessments and the eligibility process to be accessible to all countries. The final theme described issues of **diversity and difference,** which included paying attention to comorbidity, the severity and spectrum of autism, and the need to consider social and cultural differences. Table 4 displays each theme with sub-themes and example quotes. Table 4 also shows the survey questions for R2 and R3, that were developed in response to the themes and conceptual map. Areas for future research are also identified in italics.

Rounds 2 and 3

Group and individual responses from R2 were presented back to participants in R3. Questions that reached consensus in R2 were not removed from R3, as research about establishing global eligibility criteria for a diagnosis of autism in a sporting context is in a formative stage. It was therefore important to allow participants to review responses in relation to overall group feedback.

Table 4

Themes, Sub-themes, and Questions

Theme 1: High-quality Process		
Sub-themes	**Example Quotes**	**R2 & R3 survey questions based on sub-theme**
Fair process of inclusion and exclusion	[For] any eligibility process, standardising measures and setting standards makes for fair inclusion and exclusion P11 Standardising eligibility processes globally would be a big advantage in making it fair across the world in sporting competitions P25 Fair process for assessing athletes from across the world who have autism and wish to participate in international sport P3	Question about the acceptability having an eligibility process that followed the 5-step eligibility process for athletes with intellectual disabilities.
Clear and simple process	It should be as clear as possible…and to a high standard P13 Ease of assessment within Virtus (i.e., scrutiny of applications) need to be relatively quick and efficient. P14 [The process needs to] be clear, simple, not too onerous, possible for all/most countries to achieve it. P5	Question about the acceptability having an eligibility process that followed the 5-step eligibility process for athletes with intellectual disabilities.
International agreement and recognition	International agreement on empirically validated assessment measures required for ASD diagnoses P6 It would hopefully allow for reasonable adjustments to be made across the world in an internationally accepted way that would allow greater access to sporting competitions whilst remaining fair for all. P25	Question about the acceptability having an eligibility process that followed the 5-step eligibility process for athletes with intellectual disabilities. *Further research recommendation*

	Creating eligibility criteria for athletes with autism will provide a clearer pathway for them to compete internationally in a fair category. P11 Internationally recognised process and standard. P5	
Theme 2: Standardised Criteria and Evidence		
Sub-themes	**Example Quotes**	**Questions based on sub-theme**
Agreement on standardised assessments	International agreement on empirically validated assessment measures required for ASD diagnoses P6 [The process] should be to a high standard i.e., try and stick to evidence based criteria in assessments P13 Use of standardised assessment measures P11 A reliable test available across countries P17	Question about the importance of including a copy of a standardised diagnostic test from closed list of autism assessments as a mandatory requirement in a portfolio of evidence Question about how acceptable the top five autism assessments from a brief review of the literature are, to be included in a gold standard closed list of assessments.
Formal diagnostic criteria	With a global process in place this would ensure that the playing field would be level for all athletes by requiring a general list of criteria to be met by all P20 Clear and concise eligibility criteria. P8 Formal diagnostic criteria: how we diagnose autism is clearly described and defined P10	Question about the acceptability of using the criteria from either the DSM-V or the ICD-11 definition of autism.
Evidencing impact on sports performance	Showing very clearly how autism affects an athlete's performance potential in sport P14 Autism may interact with sport performance in some 'unusual' but significant ways – for example the social deficit inherent with autism may not affect performance of a skill/movement in a testing setting but may impact on a person's ability to train and interact with others P15	Question about the importance of including a measure/ description of adaptive behaviour as a mandatory requirement in a portfolio of evidence Question about the desirability of including the impact of autism on sports performance and the need for a separate Virtus category for autism, in an educational resource

	Evidencing autism and impairment in specific sports. Autism affects the ability to perform in sport very differently based on the sport and the level of autism. P21	*Further research recommendation*
Theme 3: Accessibility		
Sub-themes	**Example Quotes**	**Questions based on sub-theme**
Increased opportunities and access to competitive sport	It is also important to include [autistic] people in high performance sport, and not just as a form of 'social inclusion' as is currently the case P17 At this point, we have no sports organization to organize sports competitions for athletes with autism, and people with autism are excluded from high-performance competitions and sporting events. The [eligibility] process can change that. P16 Athletes with Autism have a space to perform sports in high level P19 Focus on increasing access to competitive sport as well as informing plans for reasonable adjustments that can be expected for athletes with autism. P25 Establishing a… process for assessing athletes from across the world who have autism and wish to participate in international sport and encourage a greater level of engagement in sport for this group. P3	Identified as a benefit of the eligibility process
Accessible to all countries	We need to balance global consistency with ease of access to tests, i.e., whatever process is adopted must be accessible to all nations. P14 A consistent and reliable process that… will not exclude athletes from countries where access to internationally recognised diagnostic assessments for autism are less easily available. P3	Question about the acceptability of a two-route process for countries that cannot access the gold standard Question about what options could be used as an alternative for athletes from countries who cannot access the gold standard list of assessments

	[The process needs to be] possible for all/most countries to achieve it. P5	
Availability of assessments	Cultural and national differences in understanding and [access to] assessments of autism P9 Differences in access to diagnostic assessments between countries… availability of standardized tests in required languages. P5 As with all psychometric assessments, culture plays a role, and having normed tests in each country may prove a challenge. P11 Differences in access to instruments across countries. P10	Question about which countries the top five autism assessments from a brief review of the literature are validation in. Question about what options could be used as an alternative for athletes from countries who cannot access the gold standard list of assessments
Theme 4: Diversity and Difference		
Sub-themes	**Example Quotes**	**Questions based on sub-theme**
Attention to comorbidity	Establishing the absence or presence of concurrent conditions (lowered intellectual function, language disorder, ADD etc.) P24 Attention for comorbidity, i.e., with motor problems (80% of the individuals with autism also experience motor problems, while there is a subset of individuals without motor problems), but also comorbidity with intellectual impairment or maybe even language impairment P2	Question about the desirability of including the similarities and differences of autism and other conditions that can be confused, in an educational resource Question about the desirability of including difference between the Virtus categories of II1, 2 & 3 (difference between categories for intellectual disabilities and autism), in an educational resource *Further research recommendation*
Spectrum and severity of autism	Autism traits/presentation will vary from person to person. There is no 'one size fits all' when it comes to autism and so any eligibility criteria will need to reflect that. P8 Ensuring the assessment is flexible and reflects the diversity of ways in which ASD can present (e.g.,	Question about the importance of including a measure/ description of adaptive behaviour as a mandatory requirement in a portfolio of evidence Question about the desirability of including the impact of autism on sports performance and the need for a

	different traits/symptoms and severity of symptoms). P15	separate Virtus category for autism, in an educational resource
	A more dimensional approach (i.e., severity of autism characteristics, adaptive functioning), thus taking into account the severity levels of different autism symptoms P2	*Further research recommendation*
Social and cultural differences	Cultural and national differences in understanding and assessment of autism P9	Question about the acceptability of a two-route process for countries that cannot access the gold standard
	Differences in levels of expertise, and even understandings of what autism is (despite diagnostic criteria) P5	Question about what options could be used as an alternative for athletes from countries who cannot access the gold standard list of assessments
	Social and cultural differences in perception of and attitude towards autism P1	*Further research recommendation*

In keeping with previous Delphi studies, percentages, IQR and medians between 0 and 10, were reported in the following format (70%, 1, 8), respectively (e.g. Vosmer et al., 2009). The R3 results are presented by the areas of eligibility laid out in the conceptual map. Each area of eligibility is divided into subheadings accompanied by a table, illustrating overall consensus and consensus between groups. In line with the research aims, degrees of consensus as well as divergence are presented in tables 5-14 under the headings agree and disagree. As not all the rating scales were asked on an agree to disagree scale, in these tables, the corresponding scale labels of acceptability, importance and desirability are provided in brackets. Differences in consensus levels between overall consensus and individual groups are highlighted in bold. Questions with high overall consensus but divergence between groups, are detailed in the text. Some indicative comments are offered from participants to illustrate the reasons for their choices. From the 39 questions in R3, 12 reached high consensus and 11 reached medium consensus overall.

Process

Process of eligibility

> *"[There is a need for] a consistent and reliable process that will be credible when subject to external scrutiny, but will not exclude athletes from countries where access to internationally recognised diagnostic assessments for autism are less easily available."*
>
> *Participant 3*

There was high consensus that a five-step eligibility process should be used for confirming a diagnosis of autism in Virtus sport (96%, 1, 10; Table 5). The five-step eligibility process is laid out in Figure 8. These five steps were based on the eligibility process used by Virtus for international sport competitions in other categories, such as athletes with intellectual disabilities (Van Biesen et al., 2021).

Figure 8

Five step eligibility process based on process used for athletes with intellectual disabilities

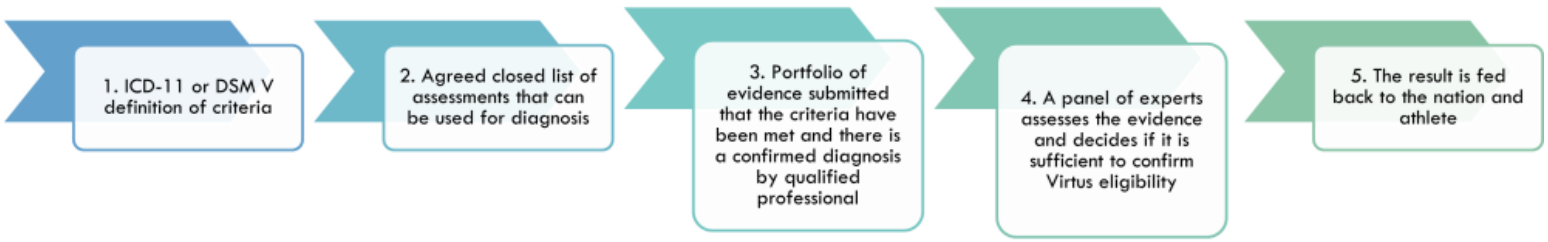


Table 5

Consensus for questions relating to the process of eligibility

		Percentages (Acceptability)		Interquartile range	Median
		Agree	Disagree		
High Consensus					
The five-step eligibility process should be used for confirming a diagnosis of autism in Virtus sport	Group 1[†]	100		1	10
	Group 2[†]	93		2	9
	Overall	96		1	10

[†] *Group 1: participants with experience administering autism assessments. Group 2: participants without experience administering autism assessments*

Definition of impairment

"Formal diagnostic criteria: how we diagnose autism needs to be clearly described and defined." Participant 10

There was high consensus (96%, 1, 10) that it would be acceptable for Virtus to use the criteria from either the DSM-V or the ICD-11 definition of autism. In R1, 83% of the

panel said that the ICD-11 and DSM-V definitions of autism were being used in countries where they were living or working (Table 6).

Table 6

Consensus for questions relating to the definition of impairment

		Percentages (Acceptability)		Interquartile range	Median
		Agree	Disagree		
High Consensus					
Virtus should use the criteria from	Group 1[†]	88.9		1	10
either the DSM-V or the ICD-11	Group 2[†]	100		2	9
definition of autism.	Overall	96		1	10

[†] Group 1: participants with experience administering autism assessments. Group 2: participants without experience administering autism assessments

Evidence

Evidence to be provided

"Standardising measures and setting standards [for evidence] makes for fair inclusion (or exclusion)." Participant 11

The expert panel was asked what needed to be included as a mandatory requirement in a portfolio of evidence. There was high consensus that a copy of a standardised diagnostic test from closed list of autism assessments (96%, 1, 10), evidence that the autism diagnosis has been certified by an appropriate professional (100%, 1, 10) and a statement from an appropriate professional that the athlete meets the DSM-V or ICD-11 criteria should be included as a mandatory requirement in a portfolio of evidence. Including a description or measure of how autism has impacted on adaptive behaviour reached medium consensus (67%, 2, 8). A brief report of developmental history reached low consensus overall (62%, 3, 8), but achieved medium consensus among participants *without* experience administering autism assessments (Table 7).

Table 7

Consensus for questions relating to what evidence should be provided

		Percentages (Importance)		Interquartile range	Median
		Agree	**Disagree**		
High Consensus					
Copy of a standardised diagnostic test from a closed list of autism assessments	Group 1†	89		1	10
	Group 2†	100		1	10
	Overall	96		1	10
Evidence that the autism diagnosis has been certified by an appropriate professional	Group 1†	100		1	10
	Group 2†	100		0	10
	Overall	100		1	10
Statement from an appropriate professional that the athlete meets DSM-V or ICD-11 criteria	Group 1†	100		1	10
	Group 2†	92		1	10
	Overall	95		1	10
Medium Consensus					
Description/ measure of how autism has impacted adaptive behaviour	Group 1†	67		2	8
	Group 2†	67		4	9
	Overall	67		2	8
Low Consensus					
Brief report of developmental history (either completed as part of an assessment or completed separately)	Group 1†	56		6	8
	Group 2†	**67**		2	8.5
	Overall	62		3	8

† *Group 1: participants with experience administering autism assessments. Group 2: participants without experience administering autism assessments*

Appropriate professionals to certify evidence

The expert panel were asked what qualifications or evidence were acceptable for a 'qualified professional' to make the initial diagnosis of autism. There was a high consensus overall that a qualified psychologist (88%, 1, 10), or a qualified psychiatrist (80%, 2, 10) were appropriate professionals to make the initial diagnosis of autism. However, there was only medium consensus for a qualified psychologist among participants *with* experience administering autism assessments (78%, 2, 10). There was only medium consensus for a qualified psychiatrist among participants *without* experience administering autism

assessments (69%, 5, 9.5). Having evidence that the diagnosis was completed by a multi-disciplinary team reached medium consensus overall (78%, 2, 9), but achieved high consensus among participants *with* experience administering autism assessments (100%, 2, 10; Table 8).

Table 8

Consensus for questions relating to appropriate professionals to certify evidence

		Percentages (Acceptability)		Interquartile range	Median
		Agree	**Disagree**		
High Consensus					
Qualified psychologist	Group 1†	**78**		2	10
	Group 2†	94		0	10
	Overall	88		1	10
Qualified psychiatrist	Group 1†	100		0	10
	Group 2†	**69**		5	9.5
	Overall	80		2	10
Medium Consensus					
Evidence diagnosis was completed by a multi-disciplinary team (MDT)	Group 1†	**100**		2	10
	Group 2†	**64**		4	9
	Overall	78		2	9
Low Consensus					
Qualified paediatrician	Group 1†	**67**		4	9
	Group 2†	**43**	29	8	7
	Overall	52		5	8
Evidence/ certificate to show that someone has been trained in a specific autism assessment	Group 1†	**75**		3	10
	Group 2†	57		5	8
	Overall	64		4	9
Lack of Consensus					
Qualified GP	Group 1†	14	43	7	5
	Group 2†	18	46	5	5
	Overall	17	45	5	5
Speech and Language therapist	Group 1†	29	29	8	5
	Group 2†	9	46	6	3
	Overall	17	39	6	4

† *Group 1: participants with experience administering autism assessments. Group 2: participants without experience administering autism assessments*

Assessments that provide acceptable evidence

The panel was asked whether a list of five autism assessments should be included in a gold standard closed list of assessments. These five assessments came from the brief literature search described in Appendix R. Participants were given information about the number of citations of the main validation paper and the countries where there was some validation for each assessment. There was high consensus to include the ADOS/ ADOS-2 (96%, 0, 10) and ADI-R (95%, 0, 10; Table 9). Including the Childhood Autism Rating Scale

Table 9

Consensus for questions relating to assessments that provide acceptable evidence

		Percentages (Acceptability)		Interquartile range	Median
		Agree	**Disagree**		
High Consensus					
Autism Diagnostic Observation Schedule/ Autism Diagnostic Observation Schedule 2nd Edition (ADOS/ ADOS-2; Lord et al., 2000)	Group 1†	100		0	10
	Group 2†	93		1	10
	Overall	96		0	10
Autism Diagnostic Interview – Revised (ADI-R; Lord, Rutter and LeCouteur 1994)	Group 1†	100		1	10
	Group 2†	92		0	10
	Overall	95		0	10
Medium Consensus					
Childhood Autism Rating Scale (CARS; Schopler et al. 1980)	Group 1†	**44**	0	5	7
	Group 2†	86		2	10
	Overall	70		3	10
Lack of Consensus					
Diagnostic Interview for Social and Communication Disorders (DISCO; Wing et al., 2002)	Group 1†	**57**		4	8
	Group 2†	42	0	4	7
	Overall	47	0	5	7
Rimland Form E1/ E2 (Rimland 1971)	Group 1†		**60**	7	1
	Group 2†	27	0	5	6
	Overall	25	19	5	5

† *Group 1: participants with experience administering autism assessments. Group 2: participants without experience administering autism assessments*

(CARS) in the gold standard list of assessments reached medium consensus overall (70%, 3, 10), but there was a lack of consensus among participants *with* experience administering autism assessments. Among participants *with* experience administering autism assessments there was low consensus for including the Diagnostic Interview for Social and Communication Disorders (DISCO) (57%, 4, 8) and low consensus for excluding Rimland Form E1/ E2 (60%, 7, 1) from the gold standard list. One participant highlighted the importance of:

> *"Agreement on empirically validated assessment measures required for ASD diagnoses." Participant 6*

Age limits for autism assessments

The panel was asked to consider whether it was acceptable to have any age limits for the autism assessments used the make the initial diagnosis. There was low consensus overall in agreement for autism assessments completed at any age (52%, 6, 8; Table 10), and for those completed with individuals aged 5 years and above (64%, 5, 8).

Table 10

Consensus for questions relating to age limits for autism assessments

		Percentages (Acceptability)		Interquartile range	Median
		Agree	**Disagree**		
Low Consensus					
5 years and up	Group 1†	**44**	44	9	6
	Group 2†	**77**		3	9
	Overall	64		5	8
Any age	Group 1†	**75**	25	7	9
	Group 2†	**39**	8	6	7
	Overall	52		6	8
18 years and up	Group 1†	0	**86**	2	1
	Group 2†	22	**44**	7	3
	Overall		63	5	2

† *Group 1: participants with experience administering autism assessments. Group 2: participants without experience administering autism assessments*

For participants *with* experience administering autism assessments having the assessment completed with any age reached medium consensus (75%, 7, 9) and there was a lack of consensus for having an age limit of 5 years and above. There was low consensus overall in disagreement with having the age limit for assessments as 18 years and above (63%, 5, 2).

Time limits for autism assessments

The panel was asked to consider how long ago an autism assessment could have been completed to provide evidence for the eligibility process. Having the assessment completed any length of time ago reached medium consensus overall (75%, 3, 9), but there was some variation in degrees of consensus between groups. There was high consensus (100%, 2, 10) among the participants *with* experience administering autism assessments and low consensus (55%, 4, 8) among panel members *without* experience administering autism assessments. Ensuring the assessment was completed recently, such as in the last 5 years, reached low consensus overall for agreement (56%, 7, 8.5) but reached low consensus in disagreement (57%, 8, 2) among panel members *with* experience administering autism assessments (Table 11).

Table 11

Consensus for questions relating to time limits for autism assessments

		Percentages (Acceptability)		Interquartile range	Median
		Agree	**Disagree**		
Medium Consensus					
Any length of time ago	Group 1[†]	**100**		2	10
	Group 2[†]	**55**		4	8
	Overall	75		3	9
Low Consensus					
Recently i.e., in the last 5 years	Group 1[†]		**57**	8	2
	Group 2[†]	**73**		3	9
	Overall	56		7	8.5

[†] Group 1: participants with experience administering autism assessments. Group 2: participants without experience administering autism assessments

Exceptions

Whether standards are universally applied

Whatever process is adopted must be accessible to all nations. Participant 14

The panel was asked how acceptable it would be to have a two-route eligibility process for countries that cannot meet the gold standard process (Appendix W). Route two would involve using an alternative to the gold standard list of assessments. Three different options were presented to the panel, described in Table 12. Having a two-route process where athletes from both routes can compete together in all competitions achieved low consensus (61%, 5, 8). There was low consensus overall (54,%, 4, 8) for having a two-route process where athletes from route two have two years of provisional eligibility, but this had a lack of consensus among panel members *with* experience administering autism assessments. There

Table 12

Consensus for questions relating to whether standards are universally applied

		Percentages (Acceptability)		Interquartile range	Median
		Agree	**Disagree**		
Low Consensus					
Athletes from both routes can compete together in all competitions	Group 1†	56		7	8
	Group 2†	64		5	8
	Overall	61		5	8
Athletes from route two have two years of provisional eligibility	Group 1†	**44**	22	7	7
	Group 2†	60		2	8
	Overall	54		4	8
Lack of Consensus					
Athletes from route two are capped at a certain level of competition	Group 1†	0	25	5	4
	Group 2†	25	50	8	2.5
	Overall	15	40	7	3

† *Group 1: participants with experience administering autism assessments. Group 2: participants without experience administering autism assessments*

was a lack of consensus about having athletes from route two capped at a certain level of competition.

"In order to be inclusive, there must be a possibility to have a 'simplified' process, however… there should be some kind of verification." Participant 2

What can be used as alternatives to gold standard assessments

It may be that athletes in some nations will not be able to be assessed using gold standard / closed list assessments. There may need to be a more narrative option to show an athlete meets the criteria. Participant 5

The expert panel was asked to consider what assessment options would be acceptable to be used as an alternative for countries who cannot access a gold standard list of assessments (Table 13). Using a combination of the WHO International Classification of

Table 13

Consensus for questions relating to alternatives to gold standard assessments

		Percentages (Acceptability)		Interquartile range	Median
		Agree	Disagree		
Medium Consensus					
Combination of ICF core sets (WHO 2003; 2007), detailed descriptive approach and validated screening measures	Group 1†	67		3	9
	Group 2†	**87**		2	9
	Overall	79		2	9
ICF core sets (WHO 2003; 2007)	Group 1†	**56**		3	8
	Group 2†	**93**		1	8
	Overall	71		1	8
Low Consensus					
Detailed descriptive approach	Group 1†	56		5	9
	Group 2†	**67**		2	8
	Overall	62		3	8
Lack of Consensus					
Validated screening measures	Group 1†	44	33	8	6
	Group 2†	**54**		4	8
	Overall	50	14	5	7.5

† *Group 1: participants with experience administering autism assessments. Group 2: participants without experience administering autism assessments*

Functioning, Disability and Health (ICF) core sets (WHO 2003; 2007), a detailed descriptive approach, and validated screening measures as an alternative reached medium consensus overall (79%, 2, 9). Just using the WHO ICF core sets as an alternative reached medium consensus overall (71%, 1, 8), but there was variation in the degrees of consensus between participants *with* experience administering autism assessments (low consensus: 56%, 3, 8) and participants *without* experience administering autism assessments (high consensus: 93%, 1, 8). There was low consensus overall about using a detailed descriptive approach (62%, 3, 8).

Other themes raised by panel

Launching an educational resource

In R1 the panel raised the idea of launching an educational resource alongside the eligibility process to help respond to some of the challenges around the themes of accessibility, diversity, and difference.

> *"[There is a need to] increase access to information about how to get an assessment, what is included, how long it takes, cost etc. Often this information is unknown, which is an initial barrier to assessment." Participant 8*

There was high consensus that Virtus should launch an educational resource that could be accessed alongside the eligibility process (92%, 0, 10). In R3 the panel was asked to consider the desirability of including the topics detailed in Table 14, as part of the educational resource. Including information about the impact of autism on sports performance achieved high consensus (88%, 2, 9). There was medium consensus overall for including the link between autism and sport (79%, 2, 9.5), the difference between Virtus eligibility categories (75%, 3, 10), similarities between autism and other diagnoses that can be confused (75%, 3, 9), and how to coach athletes with autism (70%, 3, 10) as part of the educational resource.

Table 14

Consensus for questions relating to launching an educational resource

		Percentages (Desirability)		Interquartile range	Median
		Agree	Disagree		
High Consensus					
Virtus should launch an educational resource that could be accessed alongside the eligibility process	Group 1†	89		0	10
	Group 2†	93		2	10
	Overall	92		0	10
The educational resource should cover the impact of autism on sports performance	Group 1†	89		2	9
	Group 2†	87		2	9
	Overall	88		2	9
Medium Consensus					
The educational resource should cover the link between autism and sport	Group 1†	67		4	8
	Group 2†	**87**		2	10
	Overall	79		2	9.5
The educational resource should cover the difference between Virtus categories of Intellectual Impairment 1, 2 & 3	Group 1†	**56**		3	9
	Group 2†	**87**		2	10
	Overall	75		3	10
The educational resource should cover similarities between autism and other diagnoses that could be confused	Group 1†	78		2	10
	Group 2†	73		4	9
	Overall	75		3	9
The educational resource should cover how to coach athletes with autism	Group 1†	**56**		6	8
	Group 2†	79		2	10
	Overall	70		3	10

† *Group 1: participants with experience administering autism assessments. Group 2: participants without experience administering autism assessments*

Discussion

This Delphi study gained an understanding of a panel of experts' views about the gold standard for confirming autism diagnosis in a global eligibility context for sport. This included the dimensions that are important to consider when assessing autism in a global eligibility context for sport, and the key challenges and barriers.

The results of the study showed that there was high consensus around a gold standard process of eligibility, which included the definition of the impairment and the evidence that would need to be provided. There were lower levels of consensus and agreement on having an alternative process for countries that are unable to access the gold standard. Key challenges and barriers identified included social and cultural differences, attention to co-morbidity and the heterogeneity of ASD. A dilemma was highlighted between balancing a valid and reliable process with accessibility and inclusion. The findings are discussed and linked with previous empirical and theoretical literature. Strengths, limitations, and implications for future research and practice are considered.

Levels of Consensus

Process

There was high consensus overall and between groups about the definition of impairment. The panel agreed that it was acceptable to use diagnostic criteria from either the DSM-V or ICD-11 definitions of autism. The DSM-V and the ICD-11 are the most recent versions of diagnostic criteria for ASD and are widely accepted globally (Alsayouf et al., 2020). Research has shown the specificity of the DSM-V to be high, but sensitivity has varied by clinical groups previously separate in earlier versions of the DSM e.g. Asperger's and PDD-NOS (McPartland et al., 2012).

Following the five-step eligibility process that has been used in athletes with intellectual disabilities, had strong agreement across the expert panel. This five-step process has been successfully used to determine eligibility in Virtus Regional and World Championships, and in Paralympic classification systems for athletes with intellectual disabilities (Van Biesen et al., 2021).

Evidence

The results from the Delphi panel achieved some consensus about what evidence needs to be provided for confirming autism diagnosis in a global eligibility context for sport. The panel strongly agreed that a standardised diagnostic test from a closed list of assessments would be required. In line with previous research, the ADOS-2 and the ADI-R were the most recommended diagnostic assessments to be included in this gold standard closed list, which have the largest evidence base for reliability and validity (Volker & Lopata, 2008). Whilst there was medium consensus overall to include the CARS in the gold standard list, concerns were highlighted from panel members with experience administering autism assessments. Alongside the ADOS-2 and the ADI-R, the CARS is one of the most widely used assessment scales for the screening and diagnosis of ASD (Chu et al., 2022), in a number of countries (Stewart & Lee, 2017). Whilst there is good evidence for sensitivity and internal consistency, research has shown limitations in the specificity of the CARS, and it is recommended for use alongside other confirmatory assessments (Moon et al., 2019).

There was medium consensus that the autism assessment providing evidence of the impairment could have been completed any length of time ago, however this reached high agreement among participants with experience assessing autism. Despite research consensus that ASD is a complex and heterogeneous condition (Constantino & Charman, 2016), ASD is understood to be a neurodevelopmental condition where a diagnosis stays with the individual for the lifetime (Jacobs et al., 2018). The variance in consensus around the date of assessment could be due to differing knowledge about the high resource costs in autism assessment and the feasibility of assessment. UK guidelines recommend that diagnosis of autism requires assessment by a multidisciplinary team (MDT), and multidisciplinary diagnostic assessment has a high cost per individual (Galliver et al., 2017).

There were only low levels of agreement about whether there should be age limits for assessments providing evidence of autism. There was limited agreement about having an assessment completed at any age, but this achieved medium consensus among participants with experience administering autism assessments. There is evidence that ASD can be reliably diagnosed before the age of 2, from both research and clinical samples (Sicherman et al., 2021; Zuckerman et al., 2015). A number of screening and diagnostic assessments for infants and toddlers exist, however these often utilise play-based and behavioural observation methods as part of the assessment (Thabtah & Peebles, 2019).

There was high consensus about having evidence certified by an appropriate professional, and a statement that the athlete meets the diagnostic criteria. The panel agreed that this should be included in the portfolio of evidence. As presentation of ASD varies significantly (APA, 2013), diagnosis requires building an accurate picture of the individual across settings (Galliver et al., 2017). National guidelines recommend assessment by an MDT, and in the UK NICE guidelines recommend a core team of a paediatrician, speech therapist and psychologist as good clinical practice (NICE, 2011). There was agreement that a psychologist or a psychiatrist were acceptable professionals to certify the evidence. There was medium consensus overall that there should be evidence that the diagnosis was completed by a MDT, but this reached high agreement among panel members with experience administering autism assessments. The use of MDTs in low-income countries (LICs) can be restricted due to the high price and unavailability of trained health professionals (Peiris et al., 2022). Furthermore, standards and guidelines of assessment are heavily skewed towards developed countries, and LICs are underrepresented in practice guidelines (Peiris et al., 2022).

Exceptions

The results highlighted the dilemma between having a valid and reliable process verses making the eligibility process accessible to all countries. A fundamental aspect of eligibility systems in para-sport is having valid and reliable measures of impairment (Hutchinson et al., 2020). However, accessibility and inclusion is a key foundation of para-sport (Gold & Gold, 2007). Having a two-route process, where there are alternatives to the gold standard closed list of assessments for countries that are unable to access them, reached low consensus. The panel had limited agreement that athletes from both routes could compete together in all competitions.

There were two alternatives that reached medium levels of agreement overall, for countries who cannot access a gold standard list of autism assessments. The ICF core sets are part of the globally recognized framework for defining and measuring disability and health (WHO: ICF, WHO, 2001). The ICF has previously been used in classification systems for para-sport (Van Biesen et al., 2021). There is a close taxonomic relationship between the ICF and Paralympic classification that is described by Tweedy and Vanlandewijck (2011), and this has been adopted in the IPC Classification Code (International Paralympic Committee, 2015). There were higher levels of agreement about using the ICF core sets as part of a combination of alternatives which also included a detailed descriptive approach and a validated screening measure. The heterogeneous nature of ASD makes accurate assessment challenging (Christopher & Lord, 2022). Therefore, research suggests the use of multiple assessment and confirmatory tools (Huerta & Lord, 2012), and the use of multiple informants (Möricke et al., 2016), particularly where there are concerns about the reliability and validity of available assessments.

Key Challenges and Barriers

The panel raised several themes about the key challenges and barriers that need to be considered when confirming autism diagnosis in a global eligibility context for sport. As mentioned above, the panel highlighted the need to pay attention to the spectrum of autism and the variability in severity of characteristics and symptoms (Waizbard-Bartov et al., 2021). Research has shown that ASD is a complex and heterogeneous condition (Constantino & Charman, 2016). Severity in the presentation of autism can range from individuals who require almost constant attention to meet their daily needs, to those who have jobs and are able to navigate life with little or minimal support (Whiteley et al., 2021). This is further complicated by the greater prevalence of co-occurring conditions in autistic individuals (Al-Beltagi, 2021; Mannion & Leader, 2013). Severity and comorbidity present a challenge to a fair eligibility system within the context of para-sport and highlights the need for further research into evidence-based classification. This would provide greater understanding of the relationship between autism specific challenges across various sports, and levels of severity in impairment.

Social and cultural differences were highlighted by the panel as a key challenge to a fair and consistent global eligibility process. According to the conceptual framework suggested in de Leeuw et al. (2020), subtle differences due to cultural and contextual setting can be seen in levels of expression, recognition, interpretation, and reporting of autism symptoms. Alternative explanations for autism such as cultural beliefs, supernatural explanations and beliefs about autism being ‘the western disease’ (Bhavnani et al.,2021), can lead to differences in assessment processes and varying prevalence levels between developed and LICs (Matos et al., 2022).

Educational resource

The panel suggested the idea of launching an educational resource alongside the eligibility process, to increase education around autism and address some of the key challenges that were raised. Launching an educational resource reached high consensus, and could include a simple introduction to the eligibility process in plain language for the athletes (Van Biesen et al., 2021). The need for information in accessible language has been highlighted in eligibility processes for other impairments, such as athletes with intellectual disabilities (Van Biesen et al., 2021).There was high consensus about including information about the impact of autism on sports performance within the educational resource. Currently very limited research exists in this area (Vetri & Roccella, 2020), and this presents a need for future research.

Strengths and Limitations

This study makes a unique contribution by exploring consensus views on a method for establishing global eligibility criteria for a diagnosis of autism in a sporting context. The diversity of participants in terms of experience and expertise, as well as the diversity in nationalities allowed for balanced and relevant conclusions. The methodology and recruitment of diverse participants improved validity around consensus, reducing the potential of thinking as a homogenous group with little critical evaluation (Jorm, 2015). The inclusion of parents of autistic athletes alongside professional expertise, allowed for inferences to be made about the process of eligibility for everyone rather than from one viewpoint. This helped to balance validity and reliability with access and inclusion. The use of an online survey for each round, and the aggregated presentation of the panel's previous responses, encouraged participants to respond honestly, reducing the risk of social desirability bias (De Meyrick, 2003; Surowiecki, 2004). Additionally, the recruitment of over twenty participants has been recommended to produce more stable results in Delphi studies (Jorm, 2015).

However, the use of online surveys for practical reasons in the PR and R1 meant that there was less richness in the data than if a qualitative methodology, such as focus groups, had been employed. Qualitative comments were limited by online survey responses meaning there was inadequate information to interpret differences in the levels of consensus. It is therefore, acknowledged that nuances in participant views and opinions may not be fully illustrated within the study. Further qualitative research exploring the reasons for areas with a lack consensus would be useful. Areas that achieved consensus within the Delphi methodology did not represent 'correct' answers (Hasson et al., 2000). This study identified areas and components of the eligibility process important to the participants. A further limitation was the exclusion of participants without sufficient English language skills to read and respond to a survey. Although this was a necessary practical limitation, this could have impacted the recruitment of participants from certain countries and is likely to have skewed the results towards Western developed countries where English is more widely spoken.

Further Research

Further research exploring the efficacy and feasibility of global eligibility criteria for autism is needed within different populations. For example, qualitative research exploring autistic athletes' experiences of eligibility processes in sport could provide further information about the challenges and potential solutions to accessibility and inclusion. This could also help to identify cultural and social differences. Without further research corroborating the results of this study, the practical applications and clinical implications must be considered tentatively.

Understanding the impairment-performance relationship in a given sport is important in legitimizing competition for para-sports. Currently there is very limited research exploring the impact of autism on sports performance (Vetri & Roccella, 2020). Future research in this area should be specific to how autism impacts performance during competition in a particular

sport (International Paralympic Committee, 2015), and could help to develop evidence-based classification and minimum impact criteria (Runswick et al., 2021; Van Biesen et al., 2021).

Practical Application and Clinical Implications

The results from this study highlight areas of consensus around establishing global eligibility criteria for a diagnosis of autism in a sporting context. Defining what constitutes evidence of eligible impairment is the first step to facilitate access to a new para-sport category in Virtus Regional and World Championships for autistic athletes. The results of this research have been used to put together a draft eligibility process for autistic athletes in Virtus and have been disseminated through international eligibility webinars (Appendix X). This builds upon similar research in athletes with intellectual disabilities (Van Biesen et al., 2021) and vision impairment (Runswick et al., 2021). This has the potential to reduce barriers to sport for autistic individuals at the institutional level, by changing organisational rules and policy and creating increased opportunities (Nichols et al., 2019; Obrusnikova & Miccinello, 2012). Research has shown that creating opportunities at the elite level of sport can lead to benefits in reducing barriers at other socioecological levels through the 'trickle-down' effects (Lion et al., 2022). Trickle-down effects are processes by which "people are inspired by elite sport, sports people or sports events to participate themselves" (Weed, 2009, p. 4), which have been linked to elite sport event hosting, elite sporting success, and role modelling (Potwarka & Wicker, 2020). For example, research in football has shown that elite sporting success can increase amateur participation in individual memberships, clubs, and teams (Frick & Wicker, 2016). Shifting the focus to prioritising the opportunities of sport and PA opportunities for autistic individuals, can diffuse throughout the community to influence attitudes, intentions, and behaviours (Craig et al., 2021)

One specific gain of this study is the advancement of knowledge through bringing together interdisciplinary research and practice expertise. Research on eligibility processes

can act as a focal point between disciplines such as sports science, sports psychology, neuropsychology, and clinical psychology, together with coaching expertise in different sports. This has been demonstrated in eligibility and classification research for athletes with intellectual disabilities (Van Biesen et al., 2021). Using the areas of consensus about establishing a method for confirming a diagnosis of autism in a global context increases opportunities for cross-cultural research and interventions. This has implications for global research, interventions, and activities associated with autism to meet the goal of the WHO (2013) to strengthen research and support capacities for autism in different countries.

Conclusion

Reducing barriers to sport and physical activity at the institutional level is important for autistic individuals. There are currently limited opportunities for autistic individuals in para-sport, as competitions for autistic athletes do not exist unless athletes have a comorbid intellectual impairment (International Paralympic Committee, 2016). Participation in para-sport requires proof of eligible impairment as a first step towards a global eligibility process to facilitate access to international competition. This Delphi study sought to examine whether a panel of experts can arrive at a consensus opinion on a method for establishing global eligibility criteria for a diagnosis of autism in a sporting context and if so, what the opinion is. The results of the study showed that there was high consensus around a gold standard process of eligibility. There were lower levels of consensus and agreements around whether there should be an alternative process for countries that are unable to access the gold standard. Key challenges and barriers were identified including social and cultural differences, attention to co-morbidity and the heterogeneity of ASD. Further research is needed to explore how autism impacts performance during competition in particular sports, and to develop classification and minimum impact criteria to facilitate access to international competition. There is also a need to explore the efficacy and feasibility of the results of this study, by evaluating a draft global eligibility process for a diagnosis of autism in a sporting context.

References

Al-Beltagi, M. (2021). Autism medical comorbidities. *World Journal of Clinical Pediatrics, 10*(3), 15. https://doi.org/10.5409/WJCP.V10.I3.15

Almalki, S. (2016). Integrating Quantitative and Qualitative Data in Mixed Methods Research—Challenges and Benefits. *Journal of Education and Learning*, *5*(3), 288. https://doi.org/10.5539/JEL.V5N3P288

American Psychiatric Association (APA). (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5).* Washington, DC: American Psychiatric Publishing.

Bantjes, J., & Swartz, L. (2018). Social Inclusion Through Para sport: A Critical Reflection on the Current State of Play. *Physical Medicine and Rehabilitation Clinics of North America, 29*(2), 409–416. https://doi.org/10.1016/j.pmr.2018.01.006

Beiderbeck, D., Frevel, N., von der Gracht, H. A., Schmidt, S. L., & Schweitzer, V. M. (2021). Preparing, conducting, and analyzing Delphi surveys: Cross-disciplinary practices, new directions, and advancements. MethodsX, 8, 101401.

Bogardus, R. L., Martin, R. J., Richman, A. R., & Kulas, A. S. (2019). Applying the Socio-Ecological Model to barriers to implementation of ACL injury prevention programs: A systematic review. *Journal of Sport and Health Science, 8*(1), 8–16. https://doi.org/10.1016/J.JSHS.2017.11.001

Bölte, S., de Schipper, E., Robison, J. E., Wong, V. C., Selb, M., Singhal, N., ... & Zwaigenbaum, L. (2014). Classification of Functioning and Impairment: The Development of ICF Core S ets for Autism Spectrum Disorder. *Autism Research,*

7(1), 167-172.

Booth, A., & Carroll, C. (2015). How to build up the actionable knowledge base: the role of 'best fit' framework synthesis for studies of improvement in healthcare. *BMJ Quality & Safety*, *24*(11), 700–708. https://doi.org/10.1136/BMJQS-2014-003642

Borg, S. (2016). The research journal: a tool for promoting and understanding researcher development. *Http://Dx.Doi.Org/10.1177/136216880100500204*, *5*(2), 156–177. https://doi.org/10.1177/136216880100500204

Bowles, N. (1999). The Delphi technique. *Nursing Standard (Royal College of Nursing (Great Britain) : 1987)*, *13*(45), 32–36. https://doi.org/10.7748/NS1999.07.13.45.32.C2650

Braun, V., & Clarke, V. (2006a). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706QP063OA

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J. P., Chastin, S., Chou, R., Dempsey, P. C., Dipietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., … Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine*, *54*(24), 1451–1462. https://doi.org/10.1136/BJSPORTS-2020-102955

Cantrill, J. A., Sibbald, B., & Buetow, S. (1996). The Delphi and nominal group techniques in health services research. *International Journal of Pharmacy Practice*, *4*(2), 67–74. https://doi.org/10.1111/J.2042-7174.1996.TB00844.X

Case, L., Ross, S., & Yun, J. (2020). Physical activity guideline compliance among a national sample of children with various developmental disabilities. *Disability and Health Journal*, *13*(2). https://doi.org/10.1016/j.dhjo.2019.100881

Christopher, K., & Lord, C. (2022). Best practice assessments for autism spectrum disorders in schools. *Psychology in the Schools*, *59*(7), 1454–1468. https://doi.org/10.1002/PITS.22680

Chu, J. H., Bian, F., Yan, R. Y., Li, Y., Cui, Y. H., & Li, Y. (2022). Comparison of diagnostic validity of two autism rating scales for suspected autism in a large Chinese sample. *World Journal of Clinical Cases*, *10*(4), 1206. https://doi.org/10.12998/WJCC.V10.I4.1206

Colombo-Dougovito, A. M., Blagrave, A. J., & Healy, S. (2020). A grounded theory of adoption and maintenance of physical activity among autistic adults. *Autism : The International Journal of Research and Practice*, 1362361320932444. https://doi.org/10.1177/1362361320932444

Constantino, J. N., & Charman, T. (2016). Diagnosis of autism spectrum disorder: reconciling the syndrome, its diverse origins, and variation in expression. *The Lancet. Neurology*, *15*(3), 279–291. https://doi.org/10.1016/S1474-4422(15)00151-9

Craig, D. W. (2021). Examining the effectiveness of physical activity interventions for children with autism spectrum disorders – A systematic review. *Journal of Prevention & Intervention in the Community*, *50*(1), 104-115. https://doi.org/10.1080/10852352.2021.1915939

Cunningham, G. B. (2019). Diversity and Inclusion in Sport Organizations. In *Diversity and*

Inclusion in Sport Organizations. Routledge. https://doi.org/10.4324/9780429504310

Dalkey, N., & Helmer, O. (1963). An Experimental Application of the Delphi Method to the Use of Experts. *Management Science*, *9*(3), 458–467. https://doi.org/10.1287/MNSC.9.3.458

de Meyrick, J. (2003). The Delphi method and health research. *Health Education*, *103*(1), 7–16. https://doi.org/10.1108/09654280310459112/FULL/PDF

Duquesne, V., Richard, R., Andrieu, B., & Joncheray, H. (2022). Sports experiences of elite athletes with intellectual disabilities and/or autism spectrum disorders. *Sport in Society*, 1-15. https://doi.org/10.1080/17430437.2022.2070480

Dziuk, M. A., Larson, J. C. G., Apostu, A., Mahone, E. M., Denckla, M. B., & Mostofsky, S. H. (2007). Dyspraxia in autism: association with motor, social, and communicative deficits. *Developmental Medicine & Child Neurology*, *49*(10), 734–739. https://doi.org/10.1111/J.1469-8749.2007.00734.X

Frick, B., & Wicker, P. (2016). The trickle-down effect: how elite sporting success affects amateur participation in German football. *Applied economics letters*, 23(4), 259-263.

Galliver, M., Gowling, E., Farr, W., Gain, A., & Male, I. (2017). Cost of assessing a child for possible autism spectrum disorder? An observational study of current practice in child development centres in the UK. *BMJ Paediatrics Open*, *1*(1), e000052. https://doi.org/10.1136/BMJPO-2017-000052

Gold, J. R., & Gold, M. M. (2007). Access for all: the rise of the Paralympic Games. *The Journal of the Royal Society for the Promotion of Health*, *127*(3), 133–141.

https://doi.org/10.1177/1466424007077348

Gordon, T. J. (1994). The delphi method. *Futures research methodology, 2*(3), 1-30.

Green, B., Jones, M., Hughes, D., & Williams, A. (1999). Applying the Delphi technique in a study of GPs' information requirements. *Health & Social Care in the Community*, *7*(3), 198–205. https://doi.org/10.1046/J.1365-2524.1999.00176.X

Hackett, S., Masson, H., & Phillips, S. (2006). Exploring consensus in practice with youth who are sexually abusive: Findings from a delphi study of practitioner views in the United Kingdom and the Republic of Ireland. *Child Maltreatment*, *11*(2), 146–156. https://doi.org/10.1177/1077559505285744

Hahler, E. M., & Elsabbagh, M. (2015). Autism: A Global Perspective. *Current Developmental Disorders Reports*, *2*(1), 58–64. https://doi.org/10.1007/S40474-014-0033-3/METRICS

Hardy, D. J., O'Brien, A. P., Gaskin, C. J., O'Brien, A. J., Morrison-Ngatai, E., Skews, G., Ryan, T., & McNulty, N. (2004). Practical application of the Delphi technique in a bicultural mental health nursing study in New Zealand. *Journal of Advanced Nursing*, *46*(1), 95–109. https://doi.org/10.1111/J.1365-2648.2003.02969.X

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, *32*(4), 1008–1015. https://doi.org/10.1046/J.1365-2648.2000.T01-1-01567.X

Healy, S., Nacario, A., Braithwaite, R. E., & Hopper, C. (2018). The effect of physical activity interventions on youth with autism spectrum disorder: A meta-analysis.

Autism Research, *11*(6), 818–833. https://doi.org/10.1002/aur.1955

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*. https://doi.org/10.1016/J.PSYCHRES.2020.112922

Hsu, C.-C., & Sandford, B. (2007). The Delphi Technique: Making Sense of Consensus. *Practical Assessment, Research and Evaluation*, *12*(10). https://doi.org/10.7275/PDZ9-TH90

Huang, J., Du, C., Liu, J., & Tan, G. (2020). Meta-Analysis on Intervention Effects of Physical Activities on Children and Adolescents with Autism. *International Journal of Environmental Research and Public Health*, *17*(6). https://doi.org/10.3390/IJERPH17061950

Huerta, M., & Lord, C. (2012). Diagnostic Evaluation of Autism Spectrum Disorders. *Pediatric Clinics of North America*, *59*(1), 103. https://doi.org/10.1016/J.PCL.2011.10.018

Hutchinson, M. J., Phillips, J. L. K., Mason, B. S., Goosey-Tolfrey, V. L., & Beckman, E. M. (2020). Measures of impairment applicable to the classification of Paralympic athletes competing in wheelchair sports: A systematic review of validity, reliability and associations with performance. *Https://Doi.Org/10.1080/02640414.2020.1815957*, *39*(sup1), 40–61. https://doi.org/10.1080/02640414.2020.1815957

International Paralympic Committee. (2016). *International Standard for Eligible Impairments*. www.paralympic.org

Iqbal, S., & Pipon-Young, L. (2009). Methods-The Delphi method--A guide from Susanne Iqbal and Laura Pipon-Young. *Psychologist*, *22*(7), 598.

Jachyra, P., Renwick, R., Gladstone, B., Anagnostou, E., & Gibson, B. E. (2020). Physical activity participation among adolescents with autism spectrum disorder. *Https://Doi.Org/10.1177/1362361320949344*, *25*(3), 613–626. https://doi.org/10.1177/1362361320949344

Jacobs, D., Steyaert, J., Dierickx, K., & Hens, K. (2018). Implications of an Autism Spectrum Disorder Diagnosis: An Interview Study of How Physicians Experience the Diagnosis in a Young Child. *Journal of Clinical Medicine*, *7*(10). https://doi.org/10.3390/JCM7100348

Jones, J., & Hunter, D. (1995). Consensus methods for medical and health services research. *BMJ (Clinical Research Ed.)*, *311*(7001), 376. https://doi.org/10.1136/BMJ.311.7001.376

Jones, R. A., Downing, K., Rinehart, N. J., Barnett, L. M., May, T., McGillivray, J. A., Papadopoulos, N. V., Skouteris, H., Timperio, A., & Hinkley, T. (2017). Physical activity, sedentary behavior and their correlates in children with autism spectrum disorder: A systematic review. *PLoS ONE*, *12*(2). https://doi.org/10.1371/journal.pone.0172482

Jorm, A. F. (2015). Using the Delphi expert consensus method in mental health research. *The Australian and New Zealand Journal of Psychiatry*, *49*(10), 887–897. https://doi.org/10.1177/0004867415600891

Kamp-Becker, I., Tauscher, J., Wolff, N., Küpper, C., Poustka, L., Roepke, S., ... & Stroth, S.

(2021). Is the combination of ADOS and ADI-R necessary to classify ASD? rethinking the “gold standard” in diagnosing ASD. *Frontiers in Psychiatry*, *12*, https://doi.org/https://dx.doi.org/10.3389/fpsyt.2021.727308

Ketcheson, L., Hauck, J., & Ulrich, D. (2017). The effects of an early motor skill intervention on motor skills, levels of physical activity, and socialization in young children with autism spectrum disorder: A pilot study. *Autism*, *21*(4), 481–492. https://doi.org/10.1177/1362361316650611

King, B. H., Navot, N., Bernier, R., & Webb, S. J. (2014). Update on diagnostic classification in autism. *Current Opinion in Psychiatry*, *27*(2), 105–109. https://doi.org/10.1097/YCO.0000000000000040

Lion, A., Vuillemin, A., Léon, F., Delagardelle, C., & van Hoye, A. (2022). Effect of Elite Sport on Physical Activity Practice in the General Population: A Systematic Review. *Journal of Physical Activity and Health*, *20*(1), 77–93. https://doi.org/10.1123/JPAH.2022-0123

Lord, C., & Jones, R. M. (2012). Annual research review: Re-thinking the classification of autism spectrum disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *53*(5), 490–509. https://doi.org/10.1111/j.1469-7610.2012.02547.x

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of autism and developmental disorders*, *24*(5), 659-685.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., ... & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of autism and developmental disorders*, *30*(3), 205-223.

Manicolo, O., Brotzmann, M., Hagmann-von Arx, P., Grob, A., & Weber, P. (2019). Gait in children with infantile/atypical autism: Age-dependent decrease in gait variability and associations with motor skills. *European Journal of Paediatric Neurology : EJPN : Official Journal of the European Paediatric Neurology Society*, *23*(1), 117–125. https://doi.org/10.1016/J.EJPN.2018.09.011

Mannion, A., & Leader, G. (2013). Comorbidity in autism spectrum disorder: A literature review. *Research in Autism Spectrum Disorders*, *7*(12), 1595–1616. https://doi.org/10.1016/J.RASD.2013.09.006

Marsh, C. (1998). *Exploring Data. An Introduction to Data Analysis for Social Scientists.* Cambridge: Polity Press

Matos, M. B., Bara, T. S., & Cordeiro, M. L. (2022). Autism Spectrum Disorder Diagnoses: A Comparison of Countries with Different Income Levels. *Clinical Epidemiology*, *14*, 959. https://doi.org/10.2147/CLEP.S373186

Matson, J. L., Matheis, M., Burns, C. O., Esposito, G., Venuti, P., Pisula, E., ... & Goldin, R. (2017). Examining cross-cultural differences in autism spectrum disorder: A multinational comparison from Greece, Italy, Japan, Poland, and the United States. *European Psychiatry*, *42*, 70-76.

Mcleroy, K. R., Bibeau, D., Steckler, A., & Glanz, K. (1988). An ecological perspective on

health promotion programs. *Health Education Quarterly*, *15*(4), 351–377. https://doi.org/10.1177/109019818801500401

McPartland, J. C., Reichow, B., & Volkmar, F. R. (2012). Sensitivity and specificity of proposed DSM-5 diagnostic criteria for autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(4), 368–383. https://doi.org/10.1016/J.JAAC.2012.01.007

Mohd Nordin, A., Ismail, J., & Kamal Nor, N. (2021). Motor Development in Children With Autism Spectrum Disorder. *Frontiers in Pediatrics*, *9*, 598276. https://doi.org/10.3389/FPED.2021.598276

Moon, S. J., Hwang, J. S., Shin, A. L., Kim, J. Y., Bae, S. M., Sheehy-Knight, J., & Kim, J. W. (2019). Accuracy of the Childhood Autism Rating Scale: a systematic review and meta-analysis. *Developmental Medicine and Child Neurology*, *61*(9), 1030–1038. https://doi.org/10.1111/DMCN.14246/ABSTRACT

Möricke, E., Buitelaar, J. K., & Rommelse, N. N. J. (2016). Do We Need Multiple Informants When Assessing Autistic Traits? The Degree of Report Bias on Offspring, Self, and Spouse Ratings. *Journal of Autism and Developmental Disorders*, *46*(1), 164. https://doi.org/10.1007/S10803-015-2562-Y

Must, A., Phillips, S., Curtin, C., & Bandini, L. G. (2015a). Barriers to physical activity in children with autism spectrum disorders: Relationship to physical activity and screen time. *Journal of Physical Activity & Health*, *12*(4), 529. https://doi.org/10.1123/JPAH.2013-0271

Nasa, P., Jain, R., & Juneja, D. (2021). Delphi methodology in healthcare research: How to

decide its appropriateness. *World Journal of Methodology*, *11*(4), 116. https://doi.org/10.5662/WJM.V11.I4.116

National Institute for Health and Care Excellence (2012). *Autism: recognition, referral and diagnosis of children and young people on the autism spectrum.* https://www.nice.org.uk/guidance/cg128

Nichols, C., Block, M. E., Bishop, J. C., & McIntire, B. (2019). Physical activity in young adults with autism spectrum disorder: Parental perceptions of barriers and facilitators. *Autism*, *23*(6), 1398–1407. https://doi.org/10.1177/1362361318810221

Obrusnikova, I., & Miccinello, D. L. (2012). Parent perceptions of factors influencing after-school physical activity of children with autism spectrum disorders. *Adapted Physical Activity Quarterly : APAQ*, *29*(1), 63–80. https://doi.org/10.1123/APAQ.29.1.63

Operto, F. F., Pastorino, G. M. G., Scuoppo, C., Padovano, C., Vivenzio, V., Pistola, I., Belfiore, G., Rinaldi, R., de Simone, V., & Coppola, G. (2021). Adaptive Behavior, Emotional/Behavioral Problems and Parental Stress in Children With Autism Spectrum Disorder. *Frontiers in Neuroscience*, *15*, 1615. https://doi.org/10.3389/FNINS.2021.751465/BIBTEX

Ozonoff, S., Young, G. S., Goldring, S., Greiss-Hess, L., Herrera, A. M., Steele, J., Macari, S., Hepburn, S., & Rogers, S. J. (2008). Gross motor development, movement abnormalities, and early identification of autism. *Journal of Autism and Developmental Disorders*, *38*(4), 644–656. https://doi.org/10.1007/S10803-007-0430-0

Pan, C. Y. (2011). The efficacy of an aquatic program on physical fitness and aquatic skills in

children with and without autism spectrum disorders. *Research in Autism Spectrum Disorders*, *5*(1), 657–665. https://doi.org/10.1016/J.RASD.2010.08.001

Pan, C. Y., Chu, C. H., Tsai, C. L., Sung, M. C., Huang, C. Y., & Ma, W. Y. (2017). The impacts of physical activity intervention on physical and cognitive outcomes in children with autism spectrum disorder. *Autism : The International Journal of Research and Practice*, *21*(2), 190–202. https://doi.org/10.1177/1362361316633562

Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research*, *34*(5 Pt 2), 1189. /pmc/articles/PMC1089059/?report=abstract

Peiris, H., Wickramarachchi, D. C., Samarasinghe, P., Vance, P., Dahanayake, D., Kulasekara, V., & Nadeeshani, M. (2022). Diagnosing autism in low-income countries: Clinical record-based analysis in Sri Lanka. *Autism Research : Official Journal of the International Society for Autism Research*, *15*(7), 1358–1367. https://doi.org/10.1002/AUR.2765

Potwarka, L. R., & Wicker, P. (2020). Conditions under which trickle-down effects occur: A realist synthesis approach. *Sustainability*, 13(1), 69.

Reid, H., Ridout, A. J., Tomaz, S. A., Kelly, P., & Jones, N. (2022). Benefits outweigh the risks: a consensus statement on the risks of physical activity for people living with long-term conditions. *British Journal of Sports Medicine*, *56*(8), 427–438. https://doi.org/10.1136/BJSPORTS-2021-104281

Rimland, B. (1971). The differentiation of childhood psychoses: An analysis of checklists for 2,218 psychotic children. *Journal of Autism & Childhood Schizophrenia*, *1*(2), 161–174.

Rosen, N. E., Lord, C., & Volkmar, F. R. (2021). The Diagnosis of Autism: From Kanner to DSM-III to DSM-5 and Beyond. *Journal of Autism and Developmental Disorders 2021 51:12*, *51*(12), 4253–4270. https://doi.org/10.1007/S10803-021-04904-1

Runswick, O. R., Ravensbergen, R. H. J. C., Allen, P. M., & Mann, D. L. (2021). Expert opinion on classification for footballers with vision impairment: Towards evidence-based minimum impairment criteria. *Journal of Sports Sciences*, *39*(1), 30-39.https://doi.org/10.1080/02640414.2021.1881301

Rydzewska, E., Dunn, K., & Cooper, S. A. (2021). Umbrella systematic review of systematic reviews and meta-analyses on comorbid physical conditions in people with autism spectrum disorder. *The British Journal of Psychiatry : The Journal of Mental Science*, *218*(1), 10–19. https://doi.org/10.1192/BJP.2020.167

Sala, R., Amet, L., Blagojevic-Stokic, N., Shattock, P., & Whiteley, P. (2020). Bridging the Gap Between Physical Health and Autism Spectrum Disorder. *Neuropsychiatric Disease and Treatment*, *16*, 1605. https://doi.org/10.2147/NDT.S251394

Samms-Vaughan, M., Rahbar, M. H., Dickerson, A. S., Loveland, K. A., Hessabi, M., Pearson, D. A., Bressler, J., Shakespeare-Pellington, S., Grove, M. L., Coore-Desai, C., Reece, J., & Boerwinkle, E. (2017). The diagnosis of autism and autism spectrum disorder in Low and Middle Income Countries: Experience from Jamaica. *Autism : The International Journal of Research and Practice*, *21*(5), 564. https://doi.org/10.1177/1362361317698938

Saulnier, C. A., & Klaiman, C. (2022). Assessment of adaptive behavior in autism spectrum disorder. *Psychology in the Schools*, *59*(7), 1419–1429.

https://doi.org/10.1002/PITS.22690

Scharoun, S. M., Wright, K. T., Robertson-Wilson, J. E., Fletcher, P. C., Bryden, P. J., Scharoun, S. M., Wright, K. T., Robertson-Wilson, J. E., Fletcher, P. C., & Bryden, P. J. (2017). Physical Activity in Individuals with Autism Spectrum Disorders (ASD): A Review. *Autism - Paradigms, Recent Research and Clinical Applications*. https://doi.org/10.5772/66680

Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of autism and developmental disorders*, *10*(1), 91–103.

Sefen, J. A. N., Al-Salmi, S., Shaikh, Z., AlMulhem, J. T., Rajab, E., & Fredericks, S. (2020). Beneficial Use and Potential Effectiveness of Physical Activity in Managing Autism Spectrum Disorder. *Frontiers in Behavioral Neuroscience*, *14*, 587560. https://doi.org/10.3389/FNBEH.2020.587560

Sharkey, S. B., & Sharples, A. Y. (2001). An approach to consensus building using the Delphi technique: developing a learning resource in mental health. *Nurse Education Today*, *21*(5), 398–408. https://doi.org/10.1054/NEDT.2001.0573

Sicherman, N., Charite, J., Eyal, G., Janecka, M., Loewenstein, G., Law, K., Lipkin, P. H., Marvin, A. R., & Buxbaum, J. D. (2021). Clinical signs associated with earlier diagnosis of children with autism Spectrum disorder. *BMC Pediatrics*, *21*(1), 1–14. https://doi.org/10.1186/S12887-021-02551-0/FIGURES/2

Smirni, D., Smirni, P., Carotenuto, M., Parisi, L., Quatrosi, G., & Roccella, M. (2019). Noli Me Tangere: Social Touch, Tactile Defensiveness, and Communication in

Neurodevelopmental Disorders. *Brain Sciences*, *9*(12). https://doi.org/10.3390/BRAINSCI9120368

Stanish, H., Curtin, C., Must, A., Phillips, S., Maslin, M., & Bandini, L. (2015). Enjoyment, Barriers, and Beliefs about Physical Activity among Adolescents With and Without Autism Spectrum Disorder. *Adapted Physical Activity Quarterly : APAQ*, *32*(4), 302. https://doi.org/10.1123/APAQ.2015-0038

Stanish, H. I., Curtin, C., Must, A., Phillips, S., Maslin, M., & Bandini, L. G. (2017). Physical Activity Levels, Frequency, and Type Among Adolescents with and Without Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *47*(3), 785–794. https://doi.org/10.1007/s10803-016-3001-4

Stewart, L. A., & Lee, L. C. (2017). Screening for autism spectrum disorder in low- and middle-income countries: A systematic review. *Autism : The International Journal of Research and Practice*, *21*(5), 527–539. https://doi.org/10.1177/1362361316677025

Tamminen, N., Reinikainen, J., Appelqvist-Schmidlechner, K., Borodulin, K., Mäki-Opas, T., & Solin, P. (2020). Associations of physical activity with positive mental health: A population-based study. *Mental Health and Physical Activity*, *18*, 100319. https://doi.org/10.1016/j.mhpa.2020.100319

Tan, B. W. Z., Pooley, J. A., & Speelman, C. P. (2016). A Meta-Analytic Review of the Efficacy of Physical Exercise Interventions on Cognition in Individuals with Autism Spectrum Disorder and ADHD. *Journal of Autism and Developmental Disorders*, *46*(9), 3126–3143. https://doi.org/10.1007/S10803-016-2854-X

Tarr, C. W., Rineer-Hershey, A., & Larwin, K. (2020). The Effects of Physical Exercise on

Stereotypic Behaviors in Autism: Small-n Meta-Analyses. *Focus on Autism and Other Developmental Disabilities*, *35*(1), 26–35. https://doi.org/10.1177/1088357619881220

Thabtah, F., & Peebles, D. (2019). Early Autism Screening: A Comprehensive Review. *International Journal of Environmental Research and Public Health*, *16*(18). https://doi.org/10.3390/IJERPH16183502

Thom, R. P., Palumbo, M. L., Keary, C. J., Hooker, J. M., McDougle, C. J., & Ravichandran, C. T. (2022). Prevalence and factors associated with overweight, obesity, and hypertension in a large clinical sample of adults with autism spectrum disorder. *Scientific Reports 2022 12:1*, *12*(1), 1–11. https://doi.org/10.1038/s41598-022-13365-0

Tiner, S., Cunningham, G. B., & Pittman, A. (2021). "Physical activity is beneficial to anyone, including those with ASD": Antecedents of nurses recommending physical activity for people with autism spectrum disorder. *Autism : The International Journal of Research and Practice*, *25*(2), 576–587. https://doi.org/10.1177/1362361320970082

Tse, A. C. Y. (2020). Brief Report: Impact of a Physical Exercise Intervention on Emotion Regulation and Behavioral Functioning in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *50*(11), 4191–4198. https://doi.org/10.1007/S10803-020-04418-2/METRICS

Tse, C. Y. A., Lee, H. P., Chan, K. S. K., Edgar, V. B., Wilkinson-Smith, A., & Lai, W. H. E. (2019). Examining the impact of physical activity on sleep quality and executive

functions in children with autism spectrum disorder: A randomized controlled trial. *Autism : The International Journal of Research and Practice*, *23*(7), 1699–1710. https://doi.org/10.1177/1362361318823910

Tweedy, S. M., & Vanlandewijck, Y. C. (2011). International Paralympic Committee position stand—background and scientific principles of classification in Paralympic sport. *British journal of sports medicine, 45*(4), 259-269. https://doi.org/10.1136/bjsm.2009.065060

Vallerand, R. J. (2012). Intrinsic and Extrinsic Motivation in Sport and Physical Activity: A Review and a Look at the Future. In *Handbook of Sport Psychology: Third Edition* (pp. 59–83). John Wiley and Sons. https://doi.org/10.1002/9781118270011.ch3

Van Biesen, D., Burns, J., Mactavish, J., Van De Vliet, P., & Vanlandewijck, Y. (2021). Conceptual model of sport-specific classification for para-athletes with intellectual impairment. *Journal of Sports Sciences, 39*(1), 19-29.https://doi.org/10.1080/02640414.2021.1881280

Vanderstraeten, G. G., & Oomen, A. G. M. (2010). Sports for disabled people: a general outlook. *International Journal of Rehabilitation Research, 33*(4), 283–284. https://doi.org/10.1097/MRR.0B013E32834012B1

Vetri, L., & Roccella, M. (2020). On the Playing Field to Improve: A Goal for Autism. *Medicina*, *56*(11), 1–11. https://doi.org/10.3390/MEDICINA56110585

Vivanti, G. (2020). Ask the Editor: What is the Most Appropriate Way to Talk About Individuals with a Diagnosis of Autism? *Journal of Autism and Developmental Disorders*, *50*(2), 691–693. https://doi.org/10.1007/s10803-019-04280-x

Vosmer, S., Hackett, S., & Callanan, M. (2009). “normal” and “inappropriate” childhood sexual behaviours: Findings from a delphi study of professionals in the United Kingdom. *Journal of Sexual Aggression*, *15*(3), 275–288. https://doi.org/10.1080/13552600902915984

Waizbard-Bartov, E., Ferrer, E., Young, G. S., Heath, B., Rogers, S., Wu Nordahl, C., ... & Amaral, D. G. (2021). Trajectories of autism symptom severity change during early childhood. *Journal of autism and developmental disorders*, *51*, 227-242.

Weed, M. (2009). The potential of the demonstration effect to grow and sustain participation in sport. In *A Review Paper for Sport England; Centre for Sport, Physical Education and Activity Research (SPEAR):* Canterbury Christ Church University, Canterbury, UK.

Weir, E., Allison, C., Warrier, V., & Baron-Cohen, S. (2021). Increased prevalence of non-communicable physical health conditions among autistic adults. *Autism*, *25*(3), 681–694. https://doi.org/10.1177/1362361320953652/ASSET/IMAGES/LARGE/10.1177_1362361320953652-FIG2.JPEG

Whiteley, P., Carr, K., & Shattock, P. (2021). Research, Clinical, and Sociological Aspects of Autism. *Frontiers in Psychiatry*, *12*, 560. https://doi.org/10.3389/FPSYT.2021.481546/BIBTEX

Wing, L., Leekam, S. R., Libby, S. J., Gould, J., & Larcombe, M. (2002). The diagnostic interview for social and communication disorders: Background, inter-rater reliability and clinical use. *Journal of child psychology and psychiatry*, *43*(3), 307-325.

World Health Organization. (2013). *Autism spectrum disorders and other developmental disorders: From raising awareness to building capacity*. Geneva, Switzerland: WHO Document Production Services.

World Health Organization. (2018). *Global action plan on physical activity 2018-2030: more active people for a healthier world.* World Health Organization.

World Health Organization (2007). *International classification of functioning, disability and health: Children and youth version: ICF-CY*. Geneva: World Health Organization.

World Health Organization (2003). *The ICF Checklist version 2.1a, Clinician form for International Classification of Functioning, Disability and Health.* Geneva: World Health Organization

World Health Organisation (2023, March 29). *Autism*. https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders.

Zheng, Z., Zhang, L., Li, S., Zhao, F., Wang, Y., Huang, L., Huang, J., Zou, R., Qu, Y., & Mu, D. (2017). Association among obesity, overweight and autism spectrum disorder: a systematic review and meta-analysis. *Scientific Reports 2017 7:1*, *7*(1), 1–9. https://doi.org/10.1038/s41598-017-12003-4

Zuckerman, K. E., Lindly, O. J., & Sinche, B. K. (2015). Parental concerns, provider response, and timeliness of autism spectrum disorder diagnosis. *The Journal of Pediatrics*, *166*(6), 1431-1439.e1. https://doi.org/10.1016/J.JPEDS.2015.03.007

Section C: Appendices of Supporting Material Contents

Appendix A: Template of Data Extraction Tool

	Title of Paper	Author(s) and Year	Type of review: Systematic/ Non-Systematic	Evidence for Systemaic Review	Followed PRISMA Guidelines: Y/N
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					

Categorisation of Reviews | Sytematic Reviews | Non-Sytematic Reviews | Screening Measures | Screening Psychome ...

	Systematic Review Papers							
	Title	Author(s) and Year	Type of Measures: Screening/ Diagnostic/ Both	Population	Context	Country/ Region	Short Summary	Main Findings/ Conclusions
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								

Categorisation of Reviews | Sytematic Reviews | Non-Sytematic Reviews | Screening Measures | Screening Psychome ...

	Non-Systematic Review Papers							
	Title	Author(s) and Year	Type of Measures: Screening/ Diagnostic/ Both	Population	Context	Country/ Region	Short Summary	Main Findings/ Conclusions
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								

Categorisation of Reviews | Sytematic Reviews | **Non-Sytematic Reviews** | Screening Measures | Screening Psychome ...

	Screening Measures							
	Name of Measure	Author(s) and Year of Publication	Short or Updated Versions	Review Paper identified in	Age Group	Administration Time	Administration mode	Short Summary
1								
2								
3								
4								
5								
6								
7								
8								

Categorisation of Reviews | Sytematic Reviews | Non-Sytematic Reviews | **Screening Measures** | Screening Psychome ...

	Diagnostic Measures							
	Name of Measure	Author(s) and Year of Publication	Short or Updated Versions	Review Paper identified in	Age Group	Administration Time	Administration mode	Short Summary
1								
2								
3								
4								
5								
6								
7								
8								

... | Screening Measures | Screening Psychometrics | **Diagnostic Measures** | Diagnostic Psychometrics

	Psychometric Properties of Screening Measures										
	Name of Measure	Review Paper	Primary Study(ies) (that the psychometrics have been reported from)	Internal Consistency	Reliability	Criterion Validity	Construct Validity	Content Validity	Sensitivity	Specificity	Quality of Primary Study(ies) (where evaluated by review paper)
1											
2											
3											
4											
5											
6											
7											
8											

... | Sytematic Reviews | Non-Sytematic Reviews | Screening Measures | **Screening Psychometrics** | Diagnostic Measures ...

	Psychometric Properties of Diagnostic Measures										
	Name of Measure	Review Paper	Primary Study(ies) (that the psychometrics have been reported from)	Internal Consistency	Reliability	Criterion Validity	Construct Validity	Content Validity	Sensitivity	Specificity	Quality of Primary Study(ies) (where evaluated by review paper)
1											
2											
3											
4											
5											
6											
7											
8											

... | Screening Measures | Screening Psychometrics | Diagnostic Measures | **Diagnostic Psychometrics**

Appendix B: Extract of Citation Matrix

r	Systematic Reviews									
240	Backes et al., 2014	Baghdadli et al., 2017	Falkmer et al., 2013	Hirota et al., 2018	Levy et al., 2020	Soto et al., 2015	Stuart and Lee 2017	Sun et al., 2012	Wang et al., 2020	Wigham et al., 2019
Primary Study	1	2	3	4	5	6	7	8	9	10
Absound et al. (2011)			1							
Aguiar (2005)	1									
Albores-Gallo et al. (2012)						1				
Allison, Auyeung & Baron Cohen (2012)		1								
Andersen et al. (2011)		1								
Ashwood et al. (2016)				1						1
Assumpção Jr. et al. (1999)	1									
Baduel et al. (2017)					1					
Baird et al. (2000)					1					
Baron Cohen et al. (2005)		1	1							
Baron-Cohen et al. (2001)		1								
Bastiaansen et al. (2011)		1	1							
Becker et al. (2012)	1									
Ben-Sasson and Carter (2012)						1				
Berument et al. (1999)				1						
Berument et al. (2005)			1							
Bishop & Seltzer (2012)		1								
Bishop and Norbury (2002)			1							
Bishop et al. (2017)				1						
Boelte and Poustka (2000)			1							
Boggs et al. (2006)			1							
Bolte (2012)		1								
Bölte et al. (2008)						1				
Bölte et al. (2011)				1		1				
Booth et al. (2013)		1								1
Broadbent, Galic & Stokes (2013)		1								
Brugha et al. (2012)		1		1						
Bryson et al. (2008)			1							
Bu et al. (2017)									1	
C. Huang et al. (2016)									1	
C. Li et al. (2018)									1	
Canal-Bedia et al. (2011)					1	1				
Castro-Souza (2011)	1									
Cen et al. (2017)									1	
Chandler et al. (2007)				1						
Chang et al. (2003)		1					1			

Charman et al. (2007)				1						
Chawarska et al. (2007)			1							
Chlebowski et al. (2010)			1							
Chlebowski et al. (2013)					1					
Cicchetti et al. (2008)			1							
Cohen (2003)			1							
Cohen et al. (2003)			1							
Cohen et al. (2010)			1							
Constantino et al. (2007)				1						
Cuesta-Gómez et al. (2016)					1					
De Bildt (2015)		1								
de Bildt et al. (2003)				1						
de Bildt et al. (2004)			1							
de Bildt et al. (2009)			1							
De Bildt et al. (2016)										1
Deng et al. (2017)									1	
Derks et al. (2017)										1
Dietz et al. (2006)					1					
Duarte et al. (2003)							1			
Duvekot et al. (2014)				1						
Elwin et al. (2015)		1								
Erikkson et al. (2013)		1								1
Eugenin et al. (2015)					1					
F. Wang et al. (2017)									1	
Fombonne et al. (2012)							1			
Fusar-Poli et al. (2017)										1
Gau et al. (2011)						1				
George et al. (2014)							1			
Gillberg et al. (2001)			1							
Gong et al. (2011)								1		
Gong et al. (2015)									1	
Gotham et al. (2007)			1							
Gotham et al. (2008)			1							
Gray et al. (2008)			1							
Grodberg et al. (2014)		1								1
Guan et al. (2016)									1	
Guo and Liu (2002)								1		
Guo et al. (2002)								1		
Guo et al. (2011)						1	1	1		
Hardy et al. (2015)					1					
Hedley et al. (2010)						1	1			
Heinrich et al. (2017)										1
Hill et al. (2001)			1							
Hoekstra et al. (2008)		1								

Appendix C: Joanna Briggs Institute (JBI) Checklist for Systematic Reviews and Research Synthesis

This has been removed from the electronic copy.

Appendix D: Scale for the Assessment of Narrative Review Articles (SANRA)

This has been removed from the electronic copy.

Appendix E: An Extract of Notes and Headings in the Organisation Phase

This has been removed from the electronic copy.

Appendix F: Synthesis of Psychometric Properties for Screening Measures

Name of Measure	Version	Review Paper	Internal Consistency	Reliability	Criterion Validity	Construct Validity	Content Validity	Sensitivity	Specificity
Autism Behaviour Checklist (ABC)		Backes et al., 2014	NR	Unsatisfactory Inter-rater (evidence not evaluated)	NR	NR	NR	**Mixed** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
		Hirota et al., 2018	NR	NR	NR	NR	NR	**Unsatisfactory** (unsatisfactory evidence)	**Satisfactory** (unsatisfactory evidence)
		Stuart & Lee 2017	NR	NR	NR	NR	NR	**Mixed** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
		Sun et al., 2012	**Satisfactory** (evidence not evaluated)	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
		Wang et al., 2020	NR	NR	NR	NR	NR	NR	NR
Autism Spectrum Quotient (AQ-50) Original version		Bagdadli et al., 2017	Mixed (satisfactory evidence)	Satisfactory Test-retest (satisfactory evidence)	NR	Mixed (satisfactory evidence)	NR	Mixed (Unsatisfactory evidence)	Mixed (Unsatisfactory evidence)
		Hirota et al., 2018	NR	NR	NR	NR	NR	**Satisfactory** (Unsatisfactory evidence)	**Unsatisfactory** (Unsatisfactory evidence)
		Soto et al., 2015	**Satisfactory** (evidence not evaluated)	NR	NR	NR	NR	NR	NR
		Stuart & Lee 2017	NR	NR	NR	NR	NR	NR	NR
		Wigham et al., 2019	**Mixed** (Unsatisfactory evidence)	NR	NR	NR	NR	**Satisfactory** (Unsatisfactory evidence)	**Mixed** (Unsatisfactory evidence)
Adult Social Behavior Questionnaire (ASBQ)		Bagdadli et al., 2017	Satisfactory (unsatisfactory evidence)	NR	NR	Satisfactory (unsatisfactory evidence)	NR	NR	NR

Autism Spectrum Disorder in Adults Screening questionnaire (ASDASQ)	Bagdadli et al., 2017	Satisfactory (mixed evidence)	Satisfactory (satisfactory evidence)	NR	NR	NR	Satisfactory (unsatisfactory evidence)	Satisfactory (unsatisfactory evidence)
	Stuart & Lee 2017	NR	NR	NR	NR	NR	NR	NR
Autism Symptom Interview (ASI)	Hirota et al., 2018	NR	NR	NR	NR	NR	**Satisfactory** (unsatisfactory evidence)	**Unsatisfactory** (unsatisfactory evidence)
Autism Screening Questionnaire (ASQ)	Backes et al., 2014	**Satisfactory** (evidence not evaluated)	NR	Satisfactory (evidence not evaluated)	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
	Soto et al., 2015	**Satisfactory** (evidence not evaluated)	Satisfactory Test-retest (evidence not evaluated)	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
	Stuart & Lee 2017	NR	NR	NR	NR	NR	NR	NR
Autism Spectrum Rating Scale (ASRS)	Wang et al., 2020	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
Autism Spectrum Screening Questionnaire (ASSQ)	Hirota et al., 2018	NR	NR	NR	NR	NR	**Satisfactory** (unsatisfactory evidence)	**Satisfactory** (unsatisfactory evidence)
	Soto et al., 2015	**Satisfactory** (evidence not evaluated)	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
	Sun et al., 2012	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)

	Autism Spectrum Screening Questionnaire—Mandarin Chinese Version (ASSQ-CV)	Stuart & Lee 2017	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
Autistic Traits Assessment Scale (ATA)		Backes et al., 2014	**Satisfactory** (evidence not evaluated)	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
Childhood Asperger Syndrome Test (CAST)		Hirota et al., 2018	NR	NR	NR	NR	NR	**Satisfactory** (satisfactory evidence)	**Satisfactory** (satisfactory evidence)
		Stuart & Lee 2017	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
		Wang et al., 2020	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
Child Behaviour Checklist (CBCL)		Stuart & Lee 2017	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Mixed** (evidence not evaluated)
Children's Communication Checklist (CCC)		Hirota et al., 2018	NR	NR	NR	NR	NR	**Satisfactory** (satisfactory evidence)	**Unsatisfactory** (satisfactory evidence)
Pervasive Developmental Disorder Behaviour Inventory (PDDBI)		Falkmer et al., 2013	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Unsatisfactory** (evidence not evaluated)
Social and Communication Disorders Checklist (SCDC)		Hirota et al., 2018	NR	NR	NR	NR	NR	**Satisfactory** (unsatisfactory evidence)	**Satisfactory** (unsatisfactory evidence)
		Soto et al., 2015	**Satisfactory** (evidence not evaluated)	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)

		Stuart & lee 2017	NR	NR	NR	NR	NR	NR	NR
Social Communication Questionnaire (SCQ)		Hirota et al., 2018	NR	NR	NR	NR	NR	**Satisfactory** (mixed evidence)	**Satisfactory** (mixed evidence)
		Soto et al., 2015	**Satisfactory** (evidence not evaluated)	**Satisfactory Test-rest** (evidence not evaluated)	NR	NR	NR	NR	NR
		Stuart & Lee 2017	NR	NR	NR	NR	NR	NR	NR
		Wang et al., 2020	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
		Wigham et al., 2019	NR	NR	NR	NR	NR	**Satisfactory** (mixed evidence)	**Unsatisfactory** (mixed evidence)
Sensory Reactivity in Autism Spectrum (SR-AS)		Bagdadli et al., 2017	Satisfactory (unsatisfactory evidence)	NR	NR	Satisfactory (satisfactory evidence)	Satisfactory (unsatisfactory evidence)	NR	NR
Social Responsiveness Scale (SRS)		Hirota et al., 2018	NR	NR	NR	NR	NR	**Satisfactory** (mixed evidence)	**Mixed** (mixed evidence)
		Soto et al., 2015	NR	**Satisfactory (**evidence not evaluated)	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Mixed** (evidence not evaluated)
		Stuart & Lee 2017	NR	NR	NR	NR	NR	NR	NR
		Wang et al., 2020	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
		Wigham et al., 2019	**Satisfactory** (unsatisfactory evidence)	NR	NR	NR	NR	**Satisfactory** (unsatisfactory evidence)	**Mixed** (unsatisfactory evidence)
	SRS-Adult	Bagdadli et al., 2017	Satisfactory (mixed evidence)	Satisfactory (satisfactory evidence)	NR	Satisfactory (mixed evidence)	NR	Satisfactory (unsatisfactory evidence)	Satisfactory (unsatisfactory evidence)

SRS-2	Wigham et al., 2019	**Satisfactory** (unsatisfactory evidence)	NR	NR	NR	NR	NR	NR
SRS2-AS30/ SRS2-AS11	Bagdadli et al., 2017	Satisfactory (mixed evidence)	Satisfactory (satisfactory evidence)	NR	Satisfactory (mixed evidence)	NR	Satisfactory (unsatisfactory evidence)	Satisfactory (unsatisfactory evidence)

NR: Not Recorded; Ratings in bold highlight where statistics were available in the review paper. Ratings not in bold represent ratings where the only the interpretation of the statistic was given by the review paper. Where statistics were available, internal consistency, inter-rater reliability, and test–retest reliability were rated as *satisfactory* if the Cronbach's alpha coefficient or the Kappa coefficient was ≥70 (Cicchetti, 1994). For sensitivity and specificity values ≥70% were rated as satisfactory (Furr, 2011; Glascoe 2005). Any statistics that did not reach these thresholds were rated as *unsatisfactory*. Any range of statistics that spanned above and below these thresholds, were rated as *mixed*. Where a review paper evaluated the methodological quality of the primary studies against a standardised tool: the term *satisfactory evidence* is used to describe high-quality primary studies with few concerns, *unsatisfactory evidence* is used to describe low-quality primary studies where there are two or more concerns, and *mixed evidence* is used where some of the primary studies were rated as high-quality and others were rated as low-quality. The term *evidence not evaluated* was used to describe where the methodological quality of the primary study was not evaluated against a standardised tool by the review paper.

Appendix G: Synthesis of Psychometric Properties for Diagnostic Measures

Name of Measure	Version	Review Paper	Internal Consistency	Reliability	Criterion Validity	Construct Validity	Content Validity	Sensitivity	Specificity
Adult Asperger Assessment (AAA)*		Bagdadli et al., 2017	NR	NR	NR	Mixed (unsatisfactory evidence)	NR	Satisfactory (unsatisfactory evidence)	Satisfactory (unsatisfactory evidence)
		Falkmer et al., 2013	NR	NR	NR	NR	NR	NR	NR
ADOS	ADOS 2	Falkmer et al., 2013	NR	NR	NR	NR	NR	**Mixed** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
	ADOS-G/ ADOS 2 Module 4	Bagdadli et al., 2017	Satisfactory (mixed evidence)	Satisfactory (mixed evidence)	NR	Satisfactory (mixed evidence)	NR	Satisfactory (mixed evidence)	Satisfactory (mixed evidence)
		Wigham et al., 2019	NR	NR	NR	NR	NR	**Mixed** (mixed evidence)	**Mixed** (mixed evidence)
Autism Diagnostic Interview–Revised (ADI-R)		Backes et al., 2014	**Satisfactory (**evidence not evaluated)	**Satisfactory Inter-rater** (evidence not evaluated)	Satisfactory (evidence not evaluated)	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
		Falkmer et al., 2013	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
		Sun et al., 2012	NR	**Unsatisfactory Inter-rater Mixed Test-retest** (Evidence not evaluated)	NR	NR	NR	NR	NR

	Wigham et al., 2019	NR	NR	NR	NR	NR	**Mixed** (satisfactory evidence)	**Mixed** (satisfactory evidence)
Autism Mental Status Exam (AMSE)	Bagdadli et al., 2017	NR	Satisfactory (mixed evidence)	NR	NR	NR	Satisfactory (unsatisfactory evidence)	Satisfactory (unsatisfactory evidence)
	Wigham et al., 2019	NR	NR	NR	NR	NR	**Satisfactory** (satisfactory evidence)	**Satisfactory** (satisfactory evidence)
ASDI	Falkmer et al., 2013	NR	NR	NR	NR	NR	NR	NR
Autism Spectrum Disorders-Diagnosis for Children (ASD-DC)*	Falkmer et al., 2013	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
ASDS	Falkmer et al., 2013	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
Childhood Autism Rating Scale (CARS)*	Backes et al., 2014	**Satisfactory** (evidence not evaluated)	**Satisfactory** **Test-retest** (evidence not evaluated)	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
	Falkmer et al.,2013	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
	Stuart & Lee 2017	NR	NR	NR	NR	NR	**Mixed** (evidence not evaluated)	**Mixed** (evidence not evaluated)
	Sun et al., 2012	**Satisfactory** (evidence not evaluated)	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
CASD	Falkmer et al.,2013	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
DISCO	Falkmer et al., 2013	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Unsatisfactory** (evidence not evaluated)

GADS		Falkmer et al.,2013	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
GARS		Falkmer et al.,2013	NR	NR	NR	NR	NR	**Unsatisfactory** (evidence not evaluated)	**Unsatisfactory** (evidence not evaluated)
		Stuart & Lee 2017	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
3Di		Falkmer et al., 2013	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
Ritvo Asperger and Autism Diagnostic Scale-Revised (RAADS-R)*		Bagdadli et al., 2017	Satisfactory (satisfactory evidence)	Satisfactory Test-retest (mixed evidence)	NR	Mixed (satisfactory evidence)	Satisfactory (satisfactory evidence)	Satisfactory (unsatisfactory evidence)	Satisfactory (unsatisfactory evidence)
		Falkmer et al., 2013	NR	NR	NR	NR	NR	**Satisfactory** (evidence not evaluated)	**Satisfactory** (evidence not evaluated)
		Hirota et al., 2018	NR	NR	NR	NR	NR	**Satisfactory** (unsatisfactory evidence)	**Satisfactory** (unsatisfactory evidence)
		Wigham et al. 2019	**Satisfactory** (unsatisfactory evidence)	NR	NR	NR	NR	**Satisfactory** (unsatisfactory evidence)	**Unsatisfactory** (unsatisfactory evidence)
	RAADS-14 screen	Bagdadli et al., 2017	Satisfactory (satisfactory evidence)	NR	NR	Mixed (satisfactory evidence)	Satisfactory (satisfactory evidence)	Satisfactory (unsatisfactory evidence)	Satisfactory (unsatisfactory evidence)
		Wigham et al. 2019	**Satisfactory** (unsatisfactory evidence)	NR	NR	NR	NR	**Satisfactory** (unsatisfactory evidence)	**Mixed** (unsatisfactory evidence)

NR: Not Recorded; Ratings in bold highlight where statistics were available in the review paper. Ratings not in bold represent ratings where the only the interpretation of the statistic was given by the review paper. Where statistics were available, internal consistency, inter-rater reliability, and test–retest reliability were rated as *satisfactory* if the Cronbach's alpha coefficient or the Kappa coefficient was ≥70 (Cicchetti, 1994). For sensitivity and specificity values ≥70% were rated as satisfactory (Furr, 2011; Glascoe 2005). Any statistics that did not reach these thresholds were rated as *unsatisfactory*. Any range of statistics that spanned above and below these thresholds, were rated as *mixed*. Where a review paper evaluated the methodological quality of the primary studies against a standardised tool: the term *satisfactory evidence* is used to describe high-quality primary studies with few concerns, *unsatisfactory evidence* is used to describe low-quality primary studies where there are two or more concerns, and *mixed evidence* is used where some of the primary studies were rated as high-quality and others were rated as low-quality. The term *evidence not evaluated* was used to describe where the methodological quality of the primary study was not evaluated against a standardised tool by the review paper.

Appendix H: Invite Letter to Participants

Dear X,

I would like to invite you to take part in a research project in partnership with Virtus (https://www.virtus.sport/) and Canterbury Christ Church University. This is a brief summary of the project but there is more detailed information in the participant information sheet.

What is the research project?

The aim of this research project is to develop a global eligibility process for confirming a diagnosis of autism. We are looking at this in a para-sport context. This is so that we can develop the new class of sporting competition for autistic athletes[3] without intellectual disabilities, known as II3 in Virtus (www.virtus.sport).

What is the project needed?

There are currently limited opportunities for autistic individuals in para-sport, unless athletes also have an intellectual disability. Virtus is the international federation for athletes with intellectual impairments and is committed to developing new national and international, high level, sporting opportunities for autistic athletes. To do this we have to establish a method to confirm they are eligible to compete under the classification 'autism'. Therefore, research into the best method of confirming an autism diagnosis globally is important as a first step towards creating para- sporting competitions for autistic athletes and reducing barriers to participation. This project seeks to draw upon the expertise of a range of individuals in exploring the development of an eligibility process.

What is Virtus?

Virtus is the global organisation that governs, advocates, organises and promotes elite sport for athletes with an intellectual impairment. Virtus currently governs the eligibility of athletes with an intellectual impairment (https://www.virtus.sport/applying-for-athlete-eligibility/). Virtus set up an autism competition class (II3) as a demonstration class three years ago. To advance this opportunity, a robust system of eligibility needs to be established. This is in line with the international sport's federations rules and the International Paralympic Committee.

What does the research project involve?

We are using a Delphi panel method for this research. The Delphi method is a process used to arrive at a group opinion or decision by surveying a panel of experts. If you choose to

[3] **'Identity first' (autistic people) language rather than the 'person first' (people with autism) language has been used. This decision has been made as a response to the strong preferences of autistic people who have advocated for use of this terminology (https://autisticadvocacy.org/about-asan/identity-first-language/; Vivanti, 2020).**

participate in the study, you would form part of an expert panel, invited to respond to 3 rounds of online questionnaires. Your responses would be anonymised and aggregated and shared with the group after each round. The aggregated responses will be used to inform the questions in the later rounds. Overall, it is anticipated that 3-5 hours of your time would be required, over a period of 12 months.

Next steps

We are looking for an expert panel with a broad range of nationalities and experience in a variety of areas. We need participants to meet criteria 1 below and **one** of the other inclusion criteria listed.

1. At least three years of experience supporting autistic people;
2. Involvement and experience in using tests to assess Autism (e.g. as a psychologist who administers autism assessments);
3. Involvement and experience of sports eligibility and classification (e.g. as an eligibility officer for the Paralympics);
4. At least two years' experience coaching autistic athletes;
5. Experience of being an autistic athlete;
6. Experience of being a parent of an autistic athlete;
7. Access to professional networks related to Autism.

You also need to have sufficient English language skills to read and respond to a survey, be able access a survey on an online platform and be over 18 years old.

If you are interested in participating in this research project, then please send us an email and confirm that you meet the inclusion criteria above.

There is more information in the participant information sheet that will be sent round but if you have any questions at all then do not hesitate to get in touch using the contact details below.

Best wishes,

Anna East

Lead Researcher and Trainee Clinical Psychologist, Canterbury Christ Church University
ae285@canterbury.ac.uk

Professor Jan Burns

Emeritus Professor of Clinical Psychology, Canterbury Christ Church University/Virtus Head of Eligibility
jan.burns@canterbury.ac.uk

Dr Mark Murphy

Clinical Psychologist
mark@mmpsychology.co.uk

Appendix I: Approval from Independent Research Review Panel

This has been removed from the electronic copy.

Appendix J: Approval from Ethics Panel at Canterbury Christ Church University

This has been removed from the electronic copy.

Appendix K: Participant Information Sheet

Version Number: 2

Dated: 29.10.21

Salomons Institute for Applied Psychology
One Meadow Road, Tunbridge Wells, Kent TN1 2YG
www.canterbury.ac.uk/appliedpsychology

Information about the research

Establishing global eligibility criteria for a diagnosis of autism within a sporting context

Hello. My name is Anna East and I am a trainee clinical psychologist at Canterbury Christ Church University in the United Kingdom. I would like to invite you to take part in a research study that is in partnership with Virtus. The other researchers in this project are Professor Jan Burns from Canterbury Christ Church University and Dr Mark Murphy from the National Autistic Society. Before you decide whether to take part, it is important that you understand why the research is being done and what it would involve for you.

Talk to others about the study if you wish.

What is the purpose of the study?
Research shows that Autistic individuals are less active than the general population and those with other disabilities. A substantial reason for this is lack of opportunities and access to sport. Organised sporting competitions are an important form of physical activity. Para-sport competitions are restricted entry competitions for people with an eligible impairment. However, there are currently limited opportunities for autistic individuals in para-sport. Competitions for autistic athletes do not exist unless athletes also have an intellectual impairment. One of the reasons for the lack of para-sport opportunities for autistic athletes is due to there being no established method to confirm they are eligible to compete. Therefore research, into the best method of confirming an autism diagnosis globally is important as a first step towards creating para- sporting competitions for autistic athletes and reducing barriers to participation. This research will help Virtus to develop an evidence base to determine eligibility criteria for their II3 Autism competition class. There is a broader application of this research as establishing a method for confirming a diagnosis of autism globally is further important for global research, interventions and activities associated with autism.

Aim of the research: To identify the best method for establishing global eligibility criteria for a diagnosis of autism in a sporting context. To help Virtus develop an evidence base to determine eligibility criteria for their II3 Autism competition class.

Why have I been invited?

Participants have been identified through contact from key organisations including Virtus and other relevant international organisations. Participants have been identified based on their nationality, profession, and experience. Participants who agree to participate in the study will join an expert panel which will represent a range of nationalities, professions and experience. There will be 20-25 participants on the expert panel. Participants need to meet Criteria 1 below and one of the other criteria:

1. At least three years of experience supporting autistic people;

2. Involvement and experience in using tests to assess Autism (e.g. as a psychologist who administers autism assessments);

3. Involvement and experience of sports eligibility and classification (e.g. as an eligibility officer for the Paralympics);

4. At least two years' experience coaching athletes with Autism;

5. Experience of being an athlete with Autism;

6. Experience of being a parent of an athlete with Autism;

7. Access to professional networks related to Autism.

Do I have to take part?
Taking part in the study is completely voluntary and it is up to you whether you want to take part. If you decide not to take part in the study this will not affect you in any way. You can withdraw from the study at any time without giving a reason and this will not affect you in any way. If you decide to take part, you will be able to keep this information sheet and you will be asked to sign a consent form.

How do I consent to take part?
Before you decide whether to take part, it is important that you understand why the research is being done and what it would involve for you. Read through this information sheet and talk to others if you wish. We will ask you to email back the signed consent form if you decide to take part. You can send us an electronic signature or a typed name with an email confirming you have signed the form. Once we have received the consent form, we will send you an email to let you know when we will contact you about the next step of the research project. If you have any questions at all about the information sheet or the consent form, please use the contact details at the end of this information sheet to contact Anna East or Jan Burns.

What will happen to me if I take part?
If you decide to take part, you will be a part of an expert panel and will be involved in this research for around 12 months. You will first receive an online introductory survey with questions about your demographics and previous experience. Then as part of the panel, you will receive three rounds of online surveys about 2-3 months apart. Each round of surveys will take between 30 minutes to an hour to complete. You will be able to complete each round of the survey over 3 weeks at any time that is convenient to you. Your responses from the first round of online surveys will be analysed and used to inform questions in the later surveys. For each round of the survey, you will receive two reminder emails to complete the survey. The results of this study will be fed back to Virtus to enable them to develop their eligibility criteria.

The purpose of the panel
The purpose of the panel is to assist the with developing an appropriate eligibility process for autistic athletes. This eligibility process is just for autistic athletes who do not also have

intellectual impairments. Developing an eligibility process will enable us to create a specific competition group within Virtus for autistic athletes (II3).

Expenses and payments
As the surveys will be taking place online, there are no anticipated expenses for taking part in this research. You will not receive any reimbursement for taking part in this research.

What will I be asked to do?
You will be asked to complete an online introductory survey and then three rounds of online surveys. The surveys will have information for you to read and questions for you to answer. Each round consisting of one survey will take between 30 minutes to an hour to complete. This will all be completed online. You will be able to complete each round of the panel over 3 weeks at any time that is convenient to you. There will be around 2-3 months in between each survey.

What are the possible disadvantages and risks of taking part?
There are no expected risks to participants as a result of participating in this study but it is possible that some of participants might experience distress, when thinking about and answering some of the questions about autism, or seeing the opinion of others' which may be different to theirs. If participants think this is likely to apply to them then it is recommended that they do not participate.

What are the possible benefits of taking part?
Whilst there are no immediate benefits for those people participating in the project, it is hoped that this research will help to establish a global eligibility process for confirming a diagnosis of autism. This is a step in being able to create specific sporting competitions for autistic athletes as this research will help Virtus to develop an evidence base to determine eligibility criteria for their II3 Autism competition class.

What will happen if I don't want to carry on with the study?
You are free to withdraw your consent to participate in this research project at any time without having to give a reason. To do this send an email to Anna East at a.east285@canterbury.ac.uk saying that you would like to withdraw. If you decide to withdraw from the study this will not affect you in any way. If you withdraw from the study before the first survey round, then we will delete any data collected from you. If you withdraw from the study after any of the survey rounds have taken place, then the data collected from you in previous rounds will not be deleted, as this will have been used to construct questions in the following rounds (see the table below). If you have any questions about this when you withdraw then please email Anna East to arrange a call and we will go through it with you.

Withdrawing before Survey 1	All data will be deleted
Withdrawing after Survey 1	Your data will not be able to be deleted as it will have been used to construct questions for survey 2
Withdrawing after Survey 2	Your data will not be able to be deleted as it will have been used to construct questions for survey 2 and 3
Withdrawing after Survey 3	Your data will not be able to be deleted as it will have been used to construct questions for survey 2 and 3

Concerns and Complaints
If you have a concern about any aspect of this study, you should ask to speak to me, Anna East and I will do my best to address your concerns. You can contact me by sending me an

email at a.east285@canterbury.ac.uk. If you remain dissatisfied and wish to complain formally, you can do this by contacting the Lead Supervisor of the project, Professor Jan Burns jan.burns@canterbury.ac.uk or Dr Fergal Jones, Clinical Psychology Programme Research Director, Salomons Institute for Applied Psychology fergal.jones@canterbury.ac.uk.

Will information from or about me from taking part in the study be kept confidential?
Data about you collected through the use of online surveys will be kept confidential and will be stored securely. You will be assigned a participant number and will be kept anonymous to other members of the Delphi panel. Other members of the Delphi panel will be given a broad description of what type of experts are on the panel (e.g. sports coaches or psychologists) but identifying information will be not given at any stage. Any data gathered from you in early survey rounds will be included anonymously in later survey rounds.

All information will be stored confidentially. Any identifiable information will be kept on a separate password protected file to the research data. The only time we would have to share information about you, would be if there was a risk of harm to yourself or someone else. Only authorised researchers listed at the bottom of this information sheet will have access to identifiable information. Confidential data will be analysed and used in the write up of this research. Data will be retained for 10 years in line with the Medical Research Council and then will be disposed of securely.

Participants have the right to check the accuracy of data held about them and correct any errors.

General Data Protection Regulation (GDPR)
The following categories of personal data (as defined by the General Data Protection Regulation (GDPR)) will be processed:

Name, Age, Nationality, Gender, Profession, Health Details and Contact Details.

We have identified that the public interest in processing the personal data is:

To be able to show the diversity of participants on the expert panel and to ensure a range of nationalities and professions are represented. Personal data will be used anonymously to calculate broad descriptive characteristics of the expert panel. Personal data will be stored securely on a password protected file that is separate to any of the research data.

Data can only be accessed by, or shared with:

The researchers involved in this research project: Anna East, Professor Jan Burns and Dr Mark Murphy and an examiner involved with the Doctor of Clinical Psychology.

The identified period for the retention of personal data for this project:

Personal data will be kept for the duration of the research project which is expected to be 18 months.

If you would like to obtain further information related to how your personal data is processed for this project, please contact Anna East at a.east285@canterbury.ac.uk.

You can read further information regarding how the University processes your personal data for research purposes at the following link: Research Privacy Notice - https://www.canterbury.ac.uk/university-solicitors-office/data-protection/privacy-notices/privacy-notices.aspx

What will happen to the results of the research study?

The results of the study will be written up into a doctoral thesis and submitted to Canterbury Christ Church University. The results for this study will also be written up into a paper that will be put forward for publication. All data used in the research report will be confidential and participants will not be identified in any write up. Anonymised quotes from more open-ended questions in the surveys may be used in the write up of the research. The results of this research will be used to help Virtus to develop an evidence base to determine eligibility criteria for their II3 Autism competition class.

Who is sponsoring and funding the research?
Canterbury Christ Church University

Who has reviewed the study?
This research is taking place in the UK and has received ethical approval from Canterbury Christ Church University, a UK higher education organisation.

Further information and contact details

Lead Researcher: Anna East

a.east285@canterbury.ac.uk

Salomons Institute

1 Meadow Road

Tunbridge Wells

Kent

TN1 2YG

01227 92 7166

Lead Supervisor: Professor Jan Burns

jan.burns@canterbury.ac.uk

Supervisor: Dr Mark Murphy

mark@mmpsychology.co.uk

Appendix L: Consent Form

Canterbury Christ Church University

Salomons Institute for Applied Psychology
One Meadow Road, Tunbridge Wells, Kent TN1 2YG

Version number: 1
Participant Identification number for this study:

CONSENT FORM

Title of Project: Establishing global eligibility criteria for a diagnosis of autism within a sporting context

Name of Researcher: Anna East

Please initial box

1. I confirm that I have read and understand the information sheet dated 29.10.21 (version 2) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.	☐
3. I understand that data collected during the study may be looked at by the lead researcher Anna East and supervisors Jan Burns and Mark Murphy. I give permission for these individuals to have access to my data.	☐
4. I agree that anonymous quotes from my answers to the survey questions and other anonymous data may be used in published reports of the study findings.	☐
5. I understand that this research is taking place in the UK and has received ethical approval from Canterbury Christ Church University, a UK higher education organisation.	☐

6. I agree to take part in the above study, and I agree with the purpose of the panel ☐

Name of Participant: _____ Date: _____

Signature: _____

Name of Person taking consent: __Anna East_____ Date: _____

Signature: _____

Appendix M: Participant Debrief Sheet

Version Number: 1
Dated: 05.02.23

Salomons Institute for Applied Psychology
One Meadow Road, Tunbridge Wells, Kent TN1 2YG
www.canterbury.ac.uk/appliedpsychology

Participant Debrief Sheet

Establishing global eligibility criteria for a diagnosis of autism within a sporting context

Thank you for participating as part of the expert panel in this research study. The sheet will provide you with full details of the study in which you participated.

What was the purpose of the study?

Research shows that Autistic individuals are less active than the general population and those with other disabilities. A substantial reason for this is lack of opportunities and access to sport. Organised sporting competitions are an important form of physical activity. Para-sport competitions are restricted entry competitions for people with an eligible impairment.

However, there are currently limited opportunities for autistic individuals in para-sport. Competitions for autistic athletes do not exist unless athletes also have an intellectual impairment. One of the reasons for the lack of para-sport opportunities for autistic athletes is due to there being no established method to confirm they are eligible to compete.

Therefore research, into the best method of confirming an autism diagnosis globally is important as a first step towards creating para- sporting competitions for autistic athletes and reducing barriers to participation. This research will help Virtus to develop an evidence base to determine eligibility criteria for their II3 Autism competition class. There is a broader application of this research as establishing a method for confirming a diagnosis of autism globally is further important for global research, interventions and activities associated with autism.

Aim of the research: To identify the best method for establishing global eligibility criteria for a diagnosis of autism in a sporting context. To help Virtus develop an evidence base to determine eligibility criteria for their II3 Autism competition class.

What will happen to the results of the research study?
The results of the study will be written up into a doctoral thesis and submitted to Canterbury Christ Church University. The results for this study will also be written up into a paper that will be put forward for publication. All data used in the research report will be confidential and participants will not be identified in any write up. Anonymised quotes from more open-ended questions in the surveys may be used in the write up of the research. The results of this

research are being used to help Virtus develop an evidence base to determine eligibility criteria for their II3 Autism competition class. The provisional results from the research have been fed back to Virtus and are already having a positive influence in increasing sporting opportunities for people with autism

You will be asked by email if you would like to receive a summary of the results from this research project. Unless you opt out, you will receive a summary of the results from this research study by email in June 2023.

Will information from or about me from taking part in the study be kept confidential?
Data from this research project has been kept confidential. Only authorised researchers listed at the bottom of this information sheet have had access to identifiable information which has been stored on a separate password protected database. Data will be retained for 10 years in line with the UK Medical Research Council and then will be disposed of securely.

What will happen if I want to withdraw from the study?
As stated in the participant information sheet, since the data collection and analysis has already taken place, the data collected from you in the three rounds of the Delphi panel will **not** be able to be deleted. This is because the data from each round of the Delphi panel was used to construct the survey questions in the next round. If you have any questions about this, then please email Anna East to arrange a call and we will go through it with you.

Concerns and Complaints
If you have a concern about any aspect of this study, you should ask to speak to me, and I will do my best to address your concerns. You can contact me by emailing me at ae285@canterbury.ac.uk. If you remain dissatisfied and wish to complain formally, you can do this by contacting the Lead Supervisor of the project, Professor Jan Burns jan.burns@canterbury.ac.uk or Dr Fergal Jones, Clinical Psychology Programme Research Director, Salomons Institute for Applied Psychology fergal.jones@canterbury.ac.uk.

Who has reviewed the study?
This research is taking place in the UK and has received ethical approval from Canterbury Christ Church University, a UK higher education organisation.

Further information and contact details

Lead Researcher: Anna East
Trainee Clinical Psychologist,
Canterbury Christ Church University
ae285@canterbury.ac.uk

Salomons Institute
1 Meadow Road
Tunbridge Wells
Kent
TN1 2YG
01227 92 7166

Lead Supervisor: Professor Jan Burns
Emeritus Professor of Clinical Psychology, Canterbury Christ Church University/Virtus Head of Eligibility
jan.burns@canterbury.ac.uk

Supervisor: Dr Mark Murphy
Clinical Psychologist
mark@mmpsychology.co.uk

Appendix N: Conceptual Map for guiding the Delphi Process

1) Identify where and what we think the problems might be i.e., where we have consensus and where not,
2) What the solutions might be
3) Consensus about the best solutions

Areas of eligibility:

Process

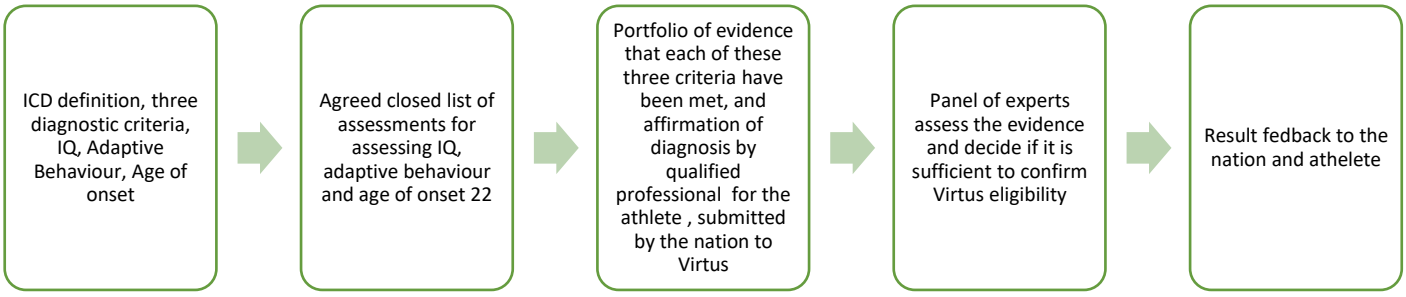
- Agreement about the process
- Agreed definition of what the impairment is

Evidence

- Agreement about what sort of evidence is acceptable
- Agreement about what assessments provide this evidence
- Agreement about how this evidence is scrutinised, and desired qualifications of the assessors

Exceptions

- Agreement about whether such standards are universally applied
- Agreement about where, why, and what exceptions are made



5 Step Eligibility Process for Athletes with Intellectual Disabilities (Van Biesen et al., 2021).

Potential ideas for each round:

Pilot Round

- Demographics
- Benefits of global eligibility process for autism
- Most important factors of global eligibility process for autism
- Main issues/ difficulties that need to be considered

Round 1

Identify where and what the problems might be:

- Definition – What definition of autism should be used
- Process – What does the eligibility process need to contain? Should the Virtus process of eligibility be used? If not, what are the problems that are highlighted?
- Evidence – what factors do autism assessments need to cover and what evidence do they need to provide?
- Should we have exceptions, or should standards be universally applied? Identifying any problems/ disagreement with this

Round 2

Identify what the solutions might be to the identified problems

- Definition – solutions to any identified problems with ICD-11 or other proposed definitions
- Process – solutions to any identified process problems in Round 1
- Portfolio of evidence – what should be included in portfolio of evidence?
- Assessments – what autism assessments provide the evidence/ factors prioritised in Round 1? How do participants rate common assessments? Where do these assessments cover? What other assessments do participants suggest?
- Exceptions – Feedback information about what the panel said about universal standards or exceptions. Solutions around making exceptions and how this should be done. Which countries might need exceptions?
- Questions about how evidence should be scrutinised – ask about problems and any solutions to these

Round 3

Consensus about the best solutions – what is the level of consensus/ agreement:

- **Level of consensus around definition.**
- **Level of consensus around process**
- **Level of consensus around portfolio of evidence**
- **Level of consensus about autism assessments**
- **Level of consensus about how evidence should be scrutinised**
- **Level of consensus about exceptions –what the exceptions are and who they might apply to**
- **Level of consensus about how evidence should be scrutinised**

Appendix O: Pilot Round Survey

This has been removed from the electronic copy.

Appendix P: Extract of Coded Transcript

This has been removed from the electronic copy.

Appendix Q: Round 1 Survey

This has been removed from the electronic copy.

Appendix R: Details of the Brief Literature Search

Brief literature search

Step 1 (Before Round 2):

Google Scholar was used to find out the number of citations of the main validation paper for a number of different autism assessments

Brief literature search to find out how roughly many countries these assessments had been validated in and what languages they had been translated into

The five autism diagnostic assessments that came out highest in these two areas combined were:

- ADOS-G/ ADOS-2
- ADI-R
- CARS
- DISCO
- Rimland Form E1/ E2

Step 2 (Before Round 3):

For the five autism assessments that came out highest in these 2 areas combined there was a more detailed literature search to list which countries these assessments had been validated in and what languages they had been translated into.

Search terms:

Search in title for the name of the measure, and then in abstract and title for the rest of the terms

For ADOS/ ADOS 2, ADI-R, Rimland Form E1/E2:

Senstiv* Specifi* Valid* Translat* Adapt* Reliabl* form* replicat* feasib* psychometric in Abstract

For DISCO and CARS:

Senstiv* Specifi* Valid* Translat* Adapt* Reliabl* form* replicat* feasib* psychometric in Abstract

Also searched Autis* in all fields (otherwise got too many results on CARS and DISCO)

Databases:

Pubmed, Embase, Medline and PsychInfo

Any Language

Publishers' websites were used to find out what languages the measures had been translated into

Add in a disclaimer that we might not have everything – these are the areas where we know there has been a validation study. *We did not look at the results of the validation studies in this brief literature search.*

Some validation means there has either been:

- A validation/ replication study
- The measure has been used in a sample of that population for a study

There is some evidence of validation in these countries from a brief literature search, but further review needs to done:

ADOS / ADOS 2

- Korean Population (South Korea)
- United Kingdom
- Germany
- Brazilian Portuguese population (published in Brazil)
- France
- South Africa (Afrikaans)
- The Netherlands
- Australia
- Sweden
- India
- Poland
- Greece
- USA (Hispanic population and rural Appalachian population)
- Austria
- Italy
- Canada
- Czech Republic
- Denmark
- Belgium
- Japan
- Norway
- Spain
- India
- Singapore
- Turkey
- China
- Russia

ADR-I

- United States (Latino population)
- France
- Sweden
- Greece
- Australia
- United Kingdom
- Italy
- Czech Republic
- Poland

- Singapore
- The Netherlands
- China + Hong Kong
- Canada
- Iran (Middle Eastern)
- Israel
- Spain
- Iceland
- Macedonia
- Finland
- Germany
- Japan
- Brazil
- South Korea
- Russia

CARS

- India
- Australia
- Jamaica
- Mexico
- Spain
- Turkey
- US
- Korea
- Turkey
- Lebanon (Lebanese-Arabic)
- Brazilian-Portuguese
- Japan
- Sweden
- Saudi Arabia
- Greece
- Brazil
- Canada
- Serbia
- Lebanon
- Italy
- Germany
- Arabic
- France
- Tanzania
- China

DISCO

- Ireland
- Australia
- UK
- US
- The Netherlands
- Sweden

RIMLAND E1-E2

- France
- US
- Australia
- Canada

Translations (From publisher's websites):

ADOS

Hungarian, English, Bulgarian, Croatian, Czech, Danish, Dutch, Finnish, French, German, Hebrew, Italian, Japanese, Korean, Mandarin: traditional characters, Norwegian, Polish, Russian, Romanian, Spanish, Swedish, Ukrainian, Simplified Chinese, Portuguese

ADI-R

English, Arabic, Danish, Dutch, Finnish, French, German, Hebrew, Hungarian, Italian, Japanese, Korean, Mandarin: traditional characters, Norwegian, Romanian, Russian, Spanish (for Spain), Swedish, Simplified Chinese, Polish, Ukrainian

CARS

English, Arabic, Bulgarian, Czech, Danish, Italian, Japanese, Korean (Chinese translated version; Pang et al., 2018)

DISCO

English

FORM E1 E2

English, Available on request: Arabic, Dutch, French, German, Hebrew, Hungarian, Italian, Japanese, Korean, Malaysian, Dutch, Polish, Portuguese, Russian, Spanish, Turkish, and Yugoslavian

Appendix S: Map Showing Validation of the Five Autism Assessments

This has been removed from the electronic copy.

Appendix T: Round 2 Survey

This has been removed from the electronic copy.

Appendix U: Example of Individualised Round 3 Survey

This has been removed from the electronic copy.

Appendix V: Excerpt from the Research and Reflections Diary

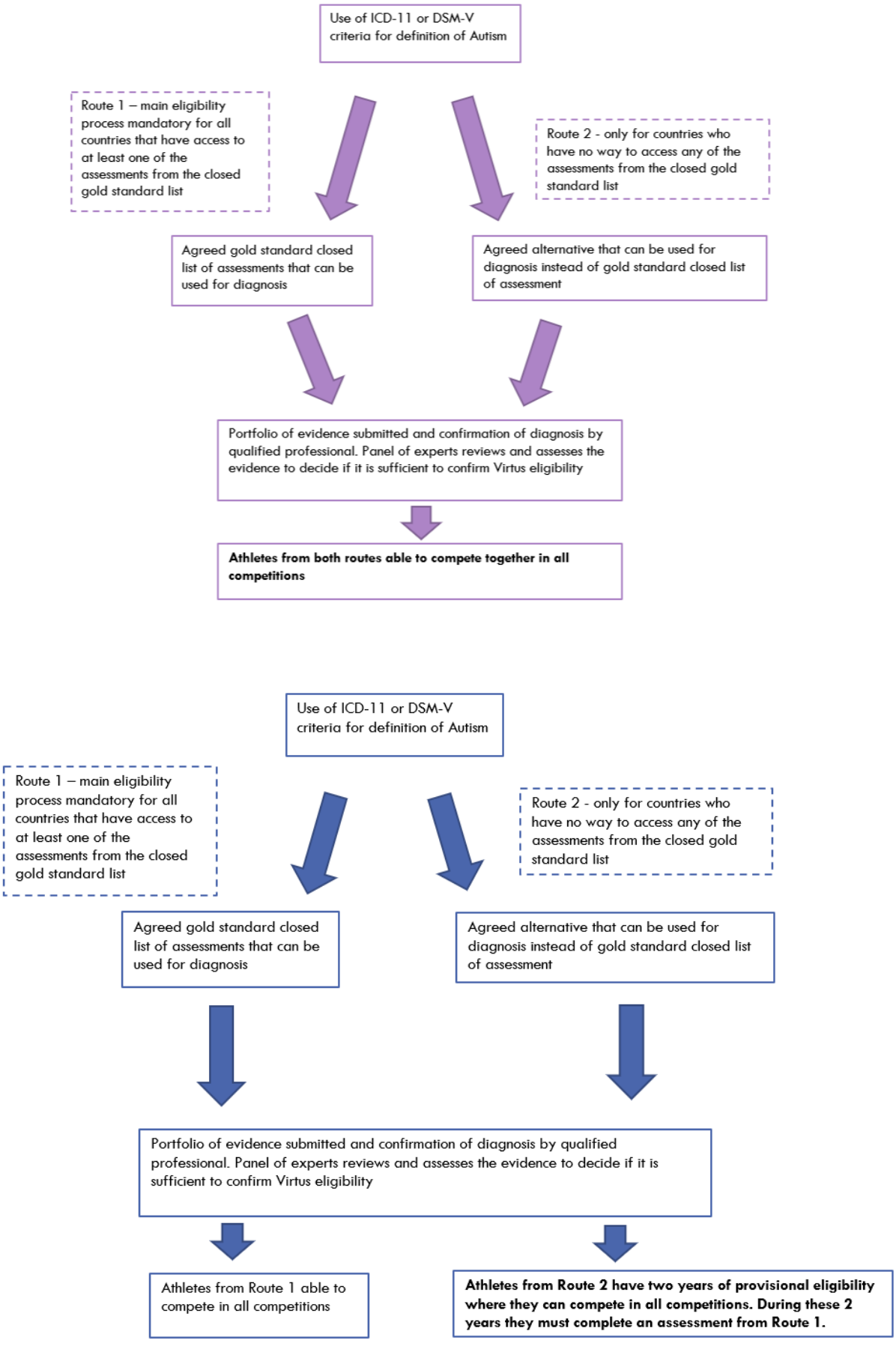
October 2021: My ethics has been approved. Going through the ethics process has made me think about how to include experts by experience in this study. My supervisor has let me know that we might be able to contact the National Autistic Society to see if there would be anyone who would be interested in getting involved as an expert by experience in the research.

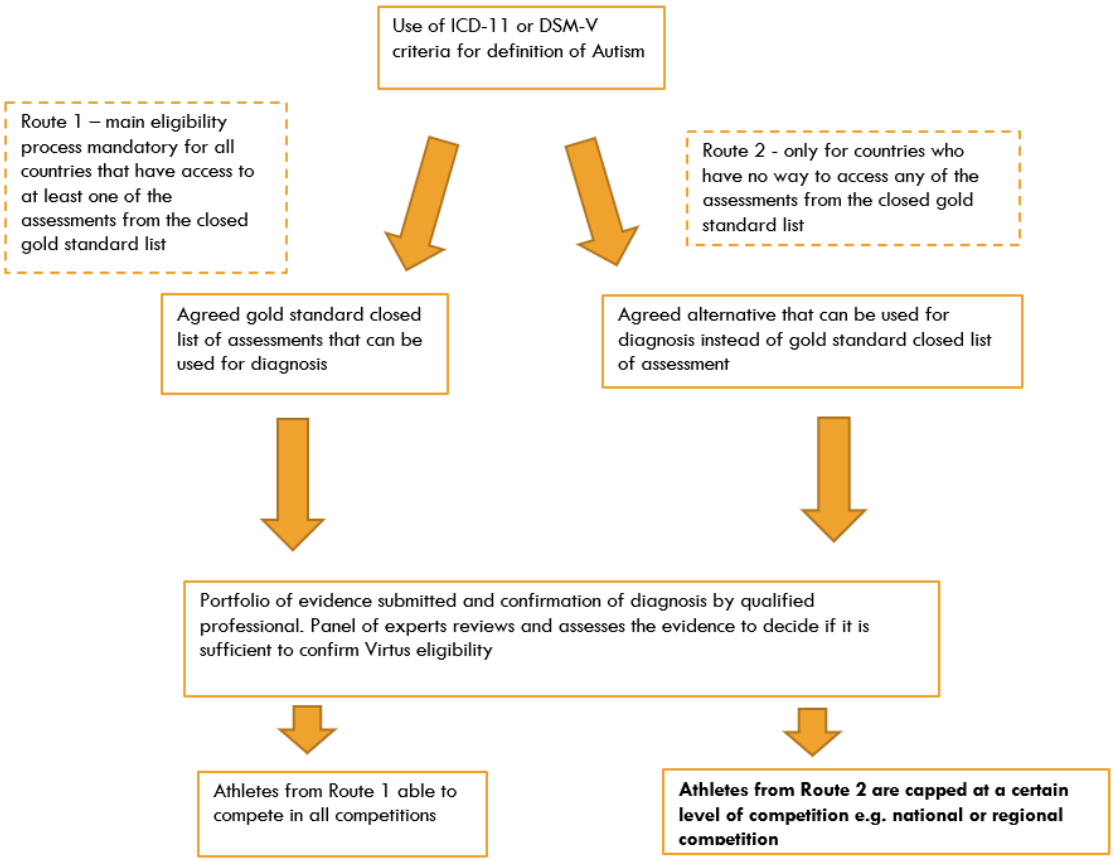
December 2021: I have been putting together a conceptual map in consultation with my research supervisor. This will help us to keep the survey rounds focussed and relevant to eligibility process for para-sport. I have been reading through the eligibility process that is used for athletes with intellectual disabilities and thinking about what the key areas of eligibility are that we might need to cover. I am curious about how an eligibility process for autism might need to be different to intellectual disabilities. It is also important to remain flexible and responsive to the responses from the expert panel. I hope that we can use this conceptual map alongside the data from each survey round when drafting the next survey.

January 2022: I am in the middle of recruiting the expert panel for the Delphi study. I am aware that it has been very hard to recruit participants from Africa and that we currently only have one participant from this continent who is South African. I was expecting it to be easier to recruit participants from Europe, but I am curious why it has been harder to recruit participants from Africa than Asia. We have had to exclude participants who cannot read and respond to a survey in English due to practical limitations. I am aware that this might have an impact on the nationalities that we have been able to recruit so far. I am excited that we have a number of parents of autistic athletes that we have recruited. We have not been able to recruit an autistic athlete to the expert panel so far.

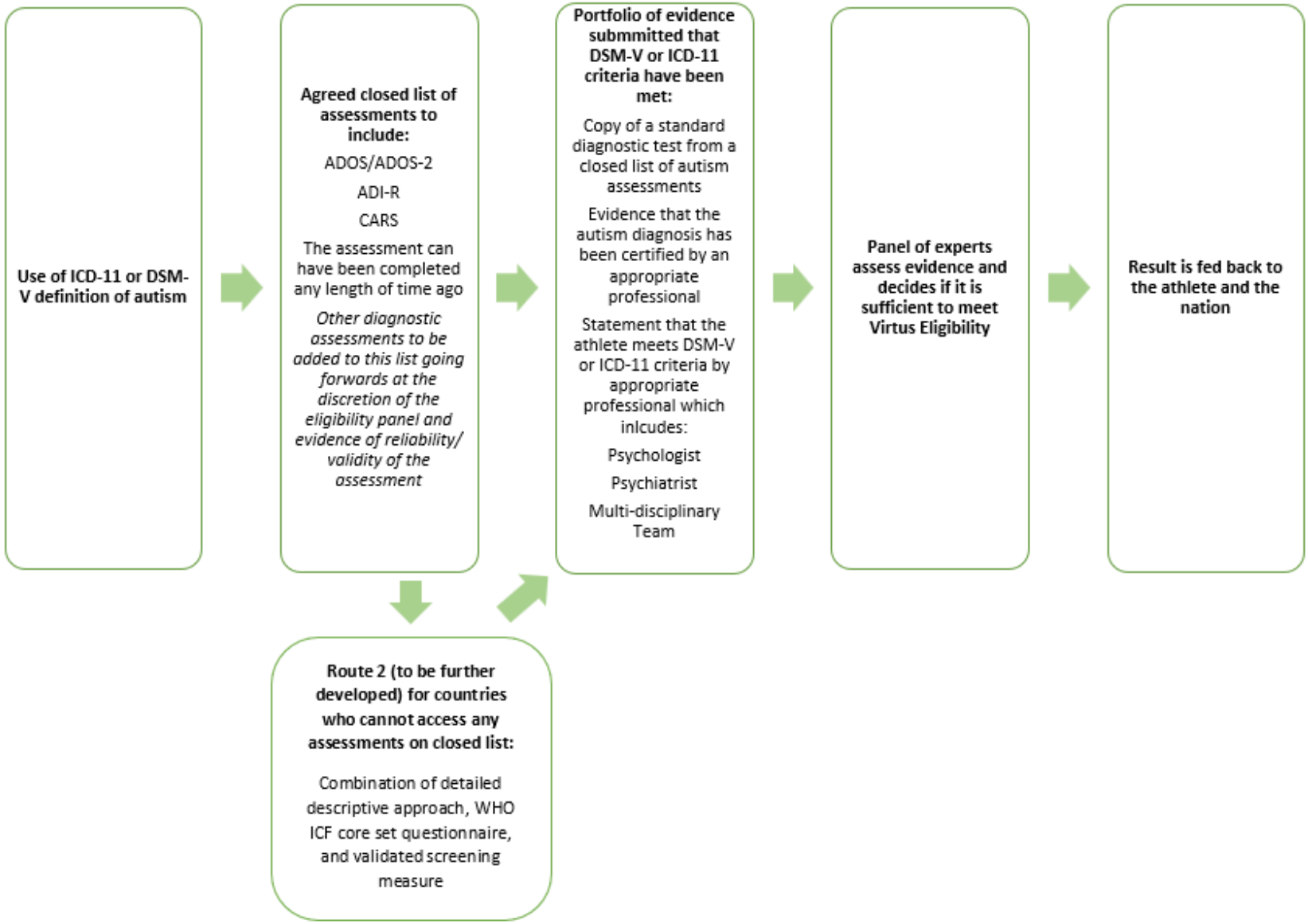
February 2022: As I am drafting the pilot round survey, I am thinking about my own experiences in sport and in assessing autism. I have competed in international sport up until the age of 18, but I am aware that I have limited knowledge and experience around para-sport. I am curious about what the experience of going through an eligibility process is like. I am also reflecting on my experience as an assistant psychologist in a CAMHS team, working with psychologists conducting autism assessments. I am aware that I have used assessments such as the ADOS-2 before, and that this feels the most familiar autism assessment to me. I am interested to know what some of the other autism assessments involve, such as what questions are in ADI-R and the CARS. I am aware that these assessments feel less familiar to me and that there are lots of other assessments for autism that I do not know about.

Appendix W: Description of Options for Two Routes from R3




Use of ICD-11 or DSM-V criteria for definition of Autism
Route 1 – main eligibility process mandatory for all countries that have access to at least one of the assessments from the closed gold standard list
Route 2 - only for countries who have no way to access any of the assessments from the closed gold standard list
Agreed gold standard closed list of assessments that can be used for diagnosis
Agreed alternative that can be used for diagnosis instead of gold standard closed list of assessment
Portfolio of evidence submitted and confirmation of diagnosis by qualified professional. Panel of experts reviews and assesses the evidence to decide if it is sufficient to confirm Virtus eligibility
Athletes from Route 1 able to compete in all competitions
Athletes from Route 2 are capped at a certain level of competition e.g. national or regional competition

Appendix X: Draft Eligibility Process for Virtus



Appendix Y: End of Study Report for Virtus

Dear members of Virtus,

Re: Increasing access to competitive sport: A Delphi study exploring the views of an expert panel on a method for establishing global eligibility criteria for a diagnosis of autism in a sporting context.

I am writing to let you know that the above research study has now been completed. I have written a summary of the project below:

Background:

There are currently limited opportunities for autistic individuals in para-sport, as competitions for autistic athletes do not exist unless athletes have a comorbid intellectual impairment. Participation in para-sport requires proof of eligible impairment as a first step towards a global eligibility process to facilitate access to international competition.

Research aims:

This study aimed to examine the views of an expert panel on a method for establishing global eligibility criteria for a diagnosis of autism in a sporting context:

1. What is the gold standard process for confirming autism diagnosis in a global eligibility context for sport?
2. What are the dimensions that are important to consider when assessing autism in a global eligibility context for sport?
3. What are the key challenges or barriers that need to be overcome?
4. What method should be used as an alternative for nations that cannot comply with the gold standard?

Method:

Twenty-seven international participants took part in a three-round Delphi panel using online surveys. This project was conducted in partnership with the organisation Virtus, the International Sports Federation for athletes with intellectual impairment. Thematic analysis of the pilot and round 1 led to the development of a round 2 online survey. A round 3 online survey was used feed-back individual and group responses and to finalise consensus.

Results:

The results of the study showed that there was high consensus around a gold standard process of eligibility, which included an agreed definition of impairment and evidence that would need to be provided. There were lower levels of consensus and agreements around whether there should be an alternative process for countries that are unable to access the gold standard. Key challenges and barriers were identified including social and cultural differences, attention to co-morbidity and the heterogeneity of ASD. Further details of the results are below:

Pilot and Round 1:

Four main themes were identified, related to the important dimensions, processes, and challenges of confirming an autism diagnosis in a global eligibility context for sport. The first was the need for a **high-quality process** of eligibility, the second theme highlighted the importance of using **standardised criteria and evidence**, the third theme identified was **accessibility** and the final theme described issues of **diversity and difference**. The survey questions for rounds 2 and 3 were developed in relation to these themes.

Rounds 2 and 3

Tables 1-4 display the areas that received the most consensus among the expert panel.

Process

Table 1. Displays the areas of process that received the most consensus among the expert panel

Areas that received high consensus overall
Using the five-step eligibility process based on the process used by Virtus for international sport competitions in other categories, such as athletes with intellectual disabilities
Using the criteria from either the DSM-V or the ICD-11 definition of autism

Evidence

Table 2. Displays the areas of evidence that received the most consensus among the expert panel

Areas that received high consensus overall
In a mandatory portfolio of evidence, the following should be provided:
Copy of a standard diagnostic test from a closed list of autism assessments

Evidence that the autism diagnosis has been certified by an appropriate professional Statement from an appropriate professional that the athlete meets DSM-V or ICD-11 criteria
That a qualified psychologist, or a qualified psychiatrist were appropriate professionals to make the initial diagnosis of autism and to certify evidence
That the Autism Diagnostic Observation Schedule (ADOS)/ Autism Diagnostic Observation Schedule 2nd Edition (ADOS-2) and the Autism Diagnostic Interview – Revised (ADI-R) should be included in a gold standard list of assessments
Areas that received medium consensus overall
In a mandatory portfolio of evidence, the following should be provided: Description/ measure of how autism has impacted adaptive behaviour
Having evidence that the initial diagnosis was completed by a multi-disciplinary team
That the Childhood Autism Rating Scale (CARS) should be included in a gold standard list of assessments
That the autism assessment can have been completed any length of time ago
Areas that received low consensus overall
In a mandatory portfolio of evidence, the following should be provided: Brief report of developmental history (either completed as part of an assessment or completed separately)
That a qualified paediatrician was an appropriate professional to make the initial diagnosis of autism and to certify evidence, or having evidence/ a certificate to show that someone has been trained in a specific autism assessment
That the autism assessment can have been completed at any age

Exceptions

Table 3. Displays the areas of exceptions that received the most consensus among the expert panel

Areas that received medium consensus overall
For countries that cannot access the assessments on the gold standard closed list: Using a combination of the WHO International Classification of Functioning, Disability and Health (ICF) core sets (WHO 2003; 2007), a detailed descriptive approach, and validated screening measures as an alternative
Areas that received low consensus overall

For countries that cannot meet the gold standard process: Having a two-route process where athletes from both routes can compete together in all competitions

Other themes raised by the panel

Table 4. Displays the areas of other themes raised by the panel that received the most consensus among the expert panel

Areas that received high consensus overall
Virtus should launch an educational resource that could be accessed alongside the eligibility process
The educational resource should cover: The impact of autism on sports performance
Areas that received medium consensus overall
The educational resource should cover: The link between autism and sport The difference between Virtus categories of Intellectual Impairment 1, 2 & 3 Similarities and differences between autism and other diagnoses that could be confused Guidance on how to coach athletes with autism

Further research is needed to explore how autism impacts performance during competition in particular sports, and to develop classification and minimum impact criteria to facilitate access to international competition.

A summary of the research project will also be sent to all participants who were part of the expert panel.

If you have any further questions regarding the project, please do not hesitate to get in contact.

Yours sincerely,

Anna East
Trainee Clinical Psychologist
Salomons Institute of Applied Psychology

Email: ae285@canterbury.ac.uk

Appendix Z: End of Study Report for Ethics

Dear ethics panel members,

Re: Increasing access to competitive sport: A Delphi study exploring the views of an expert panel on a method for establishing global eligibility criteria for a diagnosis of autism in a sporting context.

I am writing to let you know that the above research study has now been completed and submitted in partial fulfilment of the requirements of Canterbury Christ Church University Doctorate in Clinical Psychology. I have written a summary of the project below:

Background:

There are currently limited opportunities for autistic individuals in para-sport, as competitions for autistic athletes do not exist unless athletes have a comorbid intellectual impairment. Participation in para-sport requires proof of eligible impairment as a first step towards a global eligibility process to facilitate access to international competition.

Research aims:

This study aimed to examine the views of an expert panel on a method for establishing global eligibility criteria for a diagnosis of autism in a sporting context.

Method:

Twenty-seven international participants took part in a three-round Delphi panel using online surveys. This project was conducted in partnership with the organisation Virtus. Virtus is the International Sports Federation for athletes with intellectual impairment. Thematic analysis of the pilot and round 1 led to the development of a round 2 online survey. A round 3 online survey was used feed-back individual and group responses and to finalise consensus.

Results:

The results of the study showed that there was high consensus around a gold standard process of eligibility, which included an agreed definition of impairment and evidence that would need to be provided. There were lower levels of consensus and agreements around whether there should be an alternative process for countries that are unable to access the gold standard. Key challenges and barriers were identified including social and cultural differences, attention to co-morbidity and the heterogeneity of ASD. Further details of the results are below:

Pilot and Round 1:

Four main themes were identified, related to the important dimensions, processes, and challenges of confirming an autism diagnosis in a global eligibility context for sport. The first was the need for a **high-quality process** of eligibility, the second theme highlighted the importance of using **standardised criteria and evidence**, the third theme identified was **accessibility** and the final theme described issues of **diversity and difference**. The survey questions for rounds 2 and 3 were developed in relation to these themes.

Rounds 2 and 3

Tables 1-4 display the areas that received the most consensus among the expert panel.

Process

Table 1. Displays the areas of process that received the most consensus among the expert panel

Areas that received high consensus overall
Using the five-step eligibility process based on the process used by Virtus for international sport competitions in other categories, such as athletes with intellectual disabilities
Using the criteria from either the DSM-V or the ICD-11 definition of autism

Evidence

Table 2. Displays the areas of evidence that received the most consensus among the expert panel

Areas that received high consensus overall
In a mandatory portfolio of evidence, the following should be provided: Copy of a standard diagnostic test from a closed list of autism assessments Evidence that the autism diagnosis has been certified by an appropriate professional Statement from an appropriate professional that the athlete meets DSM-V or ICD-11 criteria
That a qualified psychologist, or a qualified psychiatrist were appropriate professionals to make the initial diagnosis of autism and to certify evidence

That the Autism Diagnostic Observation Schedule (ADOS)/ Autism Diagnostic Observation Schedule 2nd Edition (ADOS-2) and the Autism Diagnostic Interview – Revised (ADI-R) should be included in a gold standard list of assessments
Areas that received medium consensus overall
In a mandatory portfolio of evidence, the following should be provided: Description/ measure of how autism has impacted adaptive behaviour
Having evidence that the initial diagnosis was completed by a multi-disciplinary team
That the Childhood Autism Rating Scale (CARS) should be included in a gold standard list of assessments
That the autism assessment can have been completed any length of time ago
Areas that received low consensus overall
In a mandatory portfolio of evidence, the following should be provided: Brief report of developmental history (either completed as part of an assessment or completed separately)
That a qualified paediatrician was an appropriate professional to make the initial diagnosis of autism and to certify evidence, or having evidence/ a certificate to show that someone has been trained in a specific autism assessment
That the autism assessment can have been completed at any age

Exceptions

Table 3. Displays the areas of exceptions that received the most consensus among the expert panel

Areas that received medium consensus overall
For countries that cannot access the assessments on the gold standard closed list: Using a combination of the WHO International Classification of Functioning, Disability and Health (ICF) core sets (WHO 2003; 2007), a detailed descriptive approach, and validated screening measures as an alternative
Areas that received low consensus overall
For countries that cannot meet the gold standard process: Having a two-route process where athletes from both routes can compete together in all competitions

Other themes raised by the panel

Table 4. Displays the areas of other themes raised by the panel that received the most consensus among the expert panel

Areas that received high consensus overall
Virtus should launch an educational resource that could be accessed alongside the eligibility process
The educational resource should cover: The impact of autism on sports performance
Areas that received medium consensus overall
The educational resource should cover: The link between autism and sport The difference between Virtus categories of Intellectual Impairment 1, 2 & 3 Similarities and differences between autism and other diagnoses that could be confused Guidance on how to coach athletes with autism

Further research is needed to explore how autism impacts performance during competition in particular sports, and to develop classification and minimum impact criteria to facilitate access to international competition.

A summary of the research project will also be sent to all participants who were part of the expert panel.

If you have any further questions regarding the project, please do not hesitate to get in contact.

Yours sincerely,

Anna East
Trainee Clinical Psychologist
Salomons Institute of Applied Psychology

Email: ae285@canterbury.ac.uk

Appendix AA: End of Study Report for Participants

Dear participant,

I would like to thank you for taking part in the research on establishing a method for establishing global eligibility criteria for a diagnosis of autism in a sporting context.

Your time and contribution to this research has been very much appreciated. The study is now completed, and I am pleased to present a summary of the findings.

This study aimed to examine the views of an expert panel on a method for establishing global eligibility criteria for a diagnosis of autism in a sporting context:

5. What is the gold standard process for confirming autism diagnosis in a global eligibility context for sport?

6. What are the dimensions that are important to consider when assessing autism in a global eligibility context for sport?

7. What are the key challenges or barriers that need to be overcome?

8. What method should be used as an alternative for nations that cannot comply with the gold standard?

Twenty-seven international participants took part in the study which included academic researchers and professors, psychologists, eligibility officers, coaches, and parents of autistic athletes. The study involved completing a three-round expert panel using online surveys. This project was conducted in partnership with the organisation Virtus. Virtus is the International Sports Federation for athletes with intellectual impairment.

Results:

The results of the study showed that there was high consensus around a gold standard process of eligibility, which included an agreed definition of impairment and evidence that would need to be provided. There were lower levels of consensus and agreements around whether there should be an alternative process for countries that are unable to access the gold standard. Key challenges and barriers were identified including social and cultural differences, attention to co-morbidity and the heterogeneity of autism. Further details of the results are below:

Pilot and Round 1:

Four main themes were identified, related to the important dimensions, processes, and challenges of confirming an autism diagnosis in a global eligibility context for sport. The first was the need for a **high-quality process** of eligibility, the second theme highlighted the importance of using **standardised criteria and evidence**, the third theme identified was **accessibility** and the final theme described issues of **diversity and difference**. The survey questions for rounds 2 and 3 were developed in relation to these themes.

Rounds 2 and 3

Tables 1-4 display the areas that received the most consensus among the expert panel.

Process

Table 1. Displays the areas of process that received the most consensus among the expert panel

Areas that received high consensus overall
Using the five-step eligibility process based on the process used by Virtus for international sport competitions in other categories, such as athletes with intellectual disabilities
Using the criteria from either the DSM-V or the ICD-11 definition of autism

Evidence

Table 2. Displays the areas of evidence that received the most consensus among the expert panel

Areas that received high consensus overall
In a mandatory portfolio of evidence, the following should be provided: Copy of a standard diagnostic test from a closed list of autism assessments Evidence that the autism diagnosis has been certified by an appropriate professional Statement from an appropriate professional that the athlete meets DSM-V or ICD-11 criteria
That a qualified psychologist, or a qualified psychiatrist were appropriate professionals to make the initial diagnosis of autism and to certify evidence
That the Autism Diagnostic Observation Schedule (ADOS)/ Autism Diagnostic Observation Schedule 2nd Edition (ADOS-2) and the Autism Diagnostic Interview – Revised (ADI-R) should be included in a gold standard list of assessments
Areas that received medium consensus overall

In a mandatory portfolio of evidence, the following should be provided: Description/ measure of how autism has impacted adaptive behaviour
Having evidence that the initial diagnosis was completed by a multi-disciplinary team
That the Childhood Autism Rating Scale (CARS) should be included in a gold standard list of assessments
That the autism assessment can have been completed any length of time ago
Areas that received low consensus overall
In a mandatory portfolio of evidence, the following should be provided: Brief report of developmental history (either completed as part of an assessment or completed separately)
That a qualified paediatrician was an appropriate professional to make the initial diagnosis of autism and to certify evidence, or having evidence/ a certificate to show that someone has been trained in a specific autism assessment
That the autism assessment can have been completed at any age

Exceptions

Table 3. Displays the areas of exceptions that received the most consensus among the expert panel

Areas that received medium consensus overall
For countries that cannot access the assessments on the gold standard closed list: Using a combination of the WHO International Classification of Functioning, Disability and Health (ICF) core sets (WHO 2003; 2007), a detailed descriptive approach, and validated screening measures as an alternative
Areas that received low consensus overall
For countries that cannot meet the gold standard process: Having a two-route process where athletes from both routes can compete together in all competitions

Other themes raised by the panel

Table 4. Displays the areas of other themes raised by the panel that received the most consensus among the expert panel

Areas that received high consensus overall

Virtus should launch an educational resource that could be accessed alongside the eligibility process
The educational resource should cover: The impact of autism on sports performance
Areas that received medium consensus overall
The educational resource should cover: The link between autism and sport The difference between Virtus categories of Intellectual Impairment 1, 2 & 3 Similarities and differences between autism and other diagnoses that could be confused Guidance on how to coach athletes with autism

Further research is needed to explore how autism impacts performance during competition in particular sports. There is also a need to explore how practical and how feasible the results of this study are by evaluating a draft global eligibility process for a diagnosis of autism in a sporting context.

Further dissemination

As stated in the study information provided with the consent form, this research has been submitted as a clinical psychology doctoral thesis to Canterbury Christ Church University. This research may be published in a journal, and your anonymous quotes may be used. If you have any questions regarding the project, please do not hesitate to get in contact.

Yours sincerely,

Anna East
Trainee Clinical Psychologist
Salomons Institute of Applied Psychology

Email: ae285@canterbury.ac.uk

Appendix AB: Journal of Sports Sciences Submission Guidelines

This has been removed from the electronic copy.